STATISTICAL MODELLING OF EARTHQUAKE-INDUCED LIQUEFACTION

by

SAMSON SIM CHENG LIAO

B.S., Cornell University
(1975)

M.S., Cornell University
(1976)

Submitted to the Department of

Civil Engineering

in Partial Fulfillment of the

Requirements of the Degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1986

Signature of Author_____
Dept. of Civil Engineering, February 14, 1986

Certified by_____
Thesis Supervisor

Accepted by_____
Chairman, Departmental Committee on Graduate Students

STATISTICAL MODELLING OF EARTHQUAKE-INDUCED LIQUEFACTION

by Samson Sim Cheng Liao

Submitted to the Department of Civil Engineering on February 14, 1986
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

ABSTRACT

Statistical models are developed to calculate the conditional probability of liquefaction as a function of earthquake load and soil resistance parameters. The models are based on the analysis of a catalog (compiled in this study), which consists of 278 cases of liquefaction/non-liquefaction occurrences. Binary logistic regression (logit analysis) is the principal method used to derive the statistical models. Non-parametric kernel regression, as well as modifications of standard logistic regression, are also used. Two types of models are obtained: one uses the cyclic stress ratio (Seed and Idriss, 1971) as the earthquake load parameter; the other uses, as the load parameter, an explicit function of magnitude and distance, similar to that proposed by Davis and Berrill (1981). Both types of models use the corrected/normalized SPT $(N_1)_{60}$ value as the indicator of liquefaction resistance. Comparisons are made with other methods of liquefaction analysis.

Thesis Supervisor: Daniele Veneziano
Title: Professor of Civil Engineering

To John and Jenny,

my Parents

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my brother Jim, who made it possible for me to come to M.I.T. He assumed many of my financial and familial obligations during the past four years, for which I will remain indebted for the rest of my life.

Prof. Daniele Veneziano, the Chairman of my thesis committee, has always challenged me to take my abilities to their limits -- and then beyond. Though we have disagreed, discussed, argued -- and still disagree on many of the fine points of the research -- there is no doubt that the quality of this thesis has been enormously increased by his input.

Prof. Robert V. Whitman has played a major role in guiding the direction of this study. In addition, he has provided support and guidance in many other ways. During my years at M.I.T., Prof. Whitman has acted as a teacher, an advisor, a mentor and above all else, a friend. He is the outstanding exception that disproves the purported rule: "Nice guys always finish last".

Prof. Gregory B. Baecher has also provided me with support and advice. Interacting with him has resulted in a significant broadening of my knowledge and views on probability and statistics geotechnical engineering, and engineering in general.

There were others at M.I.T. who have assisted me in the course of this research. They include Prof. Steven R. Lerman and Dr. Joffre Swaite, who lent their time in numerous discussions on various aspects of logit analysis. Many thoughts and ideas on statistical modelling were contributed by Dr. Yusuke Honjo and Mr. Massoud Heidari. Dr.

Honjo is also credited with coding the stepwise regression procedure described in Chapter 4.

Several people from outside M.I.T. have provided data and suggestions for this study. The most noteworthy is Prof. T. Leslie Youd of Brigham Young University. Without his contributions, the data base used in the research would have been much less comprehensive and more limited in extent. Professors Raymond J. Carroll and J. Stephen Marron of the University of North Carolina at Chapel Hill are acknowledged for their advice on implementing the logistic errors in variables and kernel regression procedures.

The typing of this thesis was done by Ms. Constance Choquet and Ms. Ronni Schwartz. Mr. Steven Kennedy assisted in the production of the figures. I am grateful for their efforts and response to meeting deadlines, which were above and beyond the call of duty.

Being at M.I.T. has been like going on an expedition to a deep, dark continent. One loses track of the outside world. One loses sight of home. There are times of discovery and elation, and times of depression and drudgery. These times were shared with many others travelling similar paths, whom I wish to thank for just being friends, in particular Patrice Orleach, Mike Morrison, Mike Adams, Luc Chouinard, Chung-Tien Chin and Alec Smith. Most of all, I would like to thank my friend Ellen, soon to be my wife, for travelling with me these past few years.

Financial support for this research was provided by the National Science Foundation under Grant No. ECE 8412962.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

CHAPTER 1

INTRODUCTION

1.1 Preliminaries

1.1.1 Liquefaction, Cyclic Mobility and Liquifailure

Liquefaction is the phenomenon whereby saturated cohesionless soils (sands and silts) suffer a significant loss of shearing resistance as a result of dynamic cyclic loadings, such as those due to earthquake shaking. In some severe situations, the saturated soil may lose all shearing strength and become like a liquid ("quicksand") capable of flow -- hence the terminology. Within the geotechnical engineering profession, there are disagreements about the word "liquefaction", which was originally used by Casagrande (1936, 1938) to describe a soil with an "unstable" structure that collapses and flows under statically imposed stresses. The original definition was extended by Seed and his co-workers in the 1960's to include the general loss of shear strength in sands under dynamic loadings.

Castro (1969, 1975) has introduced the term "cyclic mobility" to describe the less severe loss of shear strength in a soil with a "stable" structure, when subjected to dynamic loading. Cyclic mobility would give rise to the development of excessive strains in a soil, but complete collapse and flow would not occur. "Liquefaction" and "cyclic mobility" are considered by Castro and his co-workers to be inherently different phenomena in terms of soil behavior mechanisms. Whitman (1985) has suggested the term "liquifailure" to encompass both flow and non-flow

types of phenomena and the terms of "liquefaction" and "cyclic mobility" for the more specific conditions.

Typically, earthquake-induced liquefaction or cyclic mobility occurs at a depth below the ground and results in some manifestation at the ground surface in the form of landslides, surface cracking, ground settlements, sand boils, distress to man-made structures or a combination of these. However, it is possible for liquefaction to occur at some depth without resulting in a surface manifestation (Ishihara, 1985).

1.1.2 A Definition

This thesis deals with the statistical analysis of case studies of actual field observations of liquefaction or non-liquefaction at sites subjected to significant earthquake shakings. Within the context of most of the case studies, it is not always practical nor possible to distinguish between true "liquefaction" and "cyclic mobility". Furthermore, one must rely on the observations (or lack thereof) of surface manifestations of liquefaction, and not conjecture about whether or not "actual" liquefaction may have occurred at depth in cases where no effects at the ground surface are seen.

Within this thesis, the use and definition of "liquefaction" is taken to be the surface manifestation of any phenomenon associated with a signficant loss of shearing resistance as a result of earthquake loadings of deposits of saturated cohesionless soils. The term "liquefaction", defined in this manner, is consistent with the term "liquifailure" proposed by Whitman (1985).

In later chapters, various statistical models are presented to evaluate the probability of liquefaction. It is important to realize that liquefaction is used here to mean that a surface manifestation concurrently occurs. This affects the way in which the statistical models are to be implemented in liquefaction risk analysis as well as the interpretations of the practical significance of the results.

1.2 Motivation

1.2.1 Uncertainties in Conventional Analysis

Several deterministic methods have been proposed to evaluate earthquake liquefaction potential. These methods range from purely empirical to highly analytical and require various degrees of laboratory testing. The prevalent approach at the present time is to use a semi-empirical chart such as that shown in Figure 1.1(a). The horizontal axis represents the strength of the soil as measured by a normalized standard penetration resistance N_1 and the vertical axis reflects the dynamic shear stress induced by an earthquake, normalized and modified by several parameters. If a point corresponding to a possible future earthquake at some site plots below the appropriate curve, it is concluded that no liquefaction would occur. Although such charts are empirical, they are founded upon a solid basis of theoretical study and laboratory investigation.

The curves in Figure 1.1(a) are based upon the study of data such as that in Figure 1.1(b), where each point represents an actual observation at a site where liquefaction did (solid dots) or did not (open circles) occur. It is evident that there is no sharp demarcation

between the two sets of observations and that several observations of liquefaction actually plot below the recommended dividing curve. This raises the question: How far should a point (representing a design situation) fall below the recommended curve to indicate an adequate margin of safety?

The problem is made more complicated by uncertainty on the actual location of the points, which results in part from lack of data (e.g., peak ground accelerations at some sites have not been measured, but are estimated from magnitude and epicentral distance), and in part from imprecise measurements (e.g., of the standard penetration resistance). There is also some disagreement about where to draw the line separating liquefaction and non-liquefaction. Various investigators using different sets of data, and relying to different degrees on laboratory results, have proposed the various lines shown in Fig. 1.2. It should be noted that there is even disagreement as to the basic shape of the line, i.e. whether it should be concave upward or concave downward.

1.2.2 Probabilistic versus Deterministic Methods

Deterministic methods of analyzing soil liquefaction give a yes/no answer as to whether liquefaction will or will not occur, or an answer in the form of a factor of safety. In either case, some consideration of probabilities must be made -- either implicitly or explicitly -- to answer questions such as: Is the risk of liquefaction high enough to justify a large monetary expenditure to improve the ground at a project site, or should the investment already made at that site be abandoned? The point here is that deterministic answers by themselves do not

generally provide guidance in decisions where uncertainties in potential failure must be weighed against potential cost.

In the context of civil engineering projects, including those involving liquefaction, the existence of risks needs to be recognized and assessed. This point was emphasized recently by Whitman (1984) and earlier by Casagrande (1965). There are some significant advantages in using probabilistic methods to assess the potential for liquefaction. The risk of liquefaction can be compared in equivalent terms to the other risks to which a structure is exposed. Uncertainties in the different inputs can be treated systematically, and those factors having the most influence on the risk can be more readily identified. Thus, probabilistic methods augment the decision process and enable a more effective transfer of experience based on subjective decisions.

1.2.3 Limitations of Previous Studies

Several investigators have tried to quantify liquefaction risk through the use of probability and statistics. Their efforts to date have been somewhat unsuccessful due to the following:

1) For probabilistic approaches, there is a lack of universally accepted physical model of liquefaction and limited understanding of all the factors that contribute to the phenomenon.
2) For statistical analyses, inappropriate procedures have been applied in some instances. For example, use has been made of discriminant analysis without recognizing that some of the assumptions underlying this method are violated in the

liquefaction case. In addition, uncertainty and bias in the data have never been modeled or accounted for in the fitting of statistical models.

A more detailed discussion of previous probabilistic and statistical approaches to liquefaction analysis is presented in Chapter 2.

1.3 Objectives and Methods

The purpose of this thesis is twofold. The first is to use appropriate statistical tools to extract information from available data on occurrences and non-occurrences of liquefaction. An important result of the analysis is the identification of earthquake loading and soil resistance variables that are statistically "significant" or "insignificant" in liquefaction analysis, thus providing insight into the underlying physical mechanisms of liquefaction.

The second purpose of the thesis is to develop simple and practical procedures for evaluating liquefaction risk. Charts similar to Fig. 1.1 are developed, but with lines corresponding to various liquefaction probability values. These procedures would be useful in site-specific evaluations of liquefaction potential and also in developing zoning maps of liquefaction hazard.

This thesis presents the use and development of a class of statistical methods applied to the analysis of historical liquefaction data. In contrast to previously used techniques, these methods

(a) are mathematically appropriate for the quantification of liquefaction probability,

(b) allow one to compare alternative selections of earthquake

intensity measures and soil resistance parameters, and

(c) can explicitly incorporate uncertainty and bias in the data.

One of these methods, known as logit analysis or logistic regression, has been applied in other disciplines, notably biological assay, medicine, econometrics, and transportation demand analysis.

Logit analysis is a particular form of a larger class of methods known as dichotomous regression, where the response variable is "binary", e.g., yes/no, success/failure, death/survival, or in the present case, liquefaction/no liquefaction. However, a limitation of logistic regression is that a certain mathematical (parametric) form is assumed in analysis. This is analogous to the assumption of a straight line relationship between the two variables in a simple least-squares linear regression. Nonparametric methods, which do not require such assumptions, were developed as part of the research and also applied to the liquefaction data.

1.4 Scope and Outline of the Thesis

Chapter 2 of this thesis presents a review of previous research and applications of probability and statistics in liquefaction analysis. The review is conducted within the framework of seismic risk analysis, which is the natural end product of such probabilistic or statistical studies.

Chapter 3 describes the catalog of liquefaction case studies which was compiled for this research. The methodology of the data compilation is outlined in detail to elucidate the problems encountered and judgments made in the course of that compilation. The limitations of the data base

are important details that must be considered in the interpretation of any results obtained from analysis.

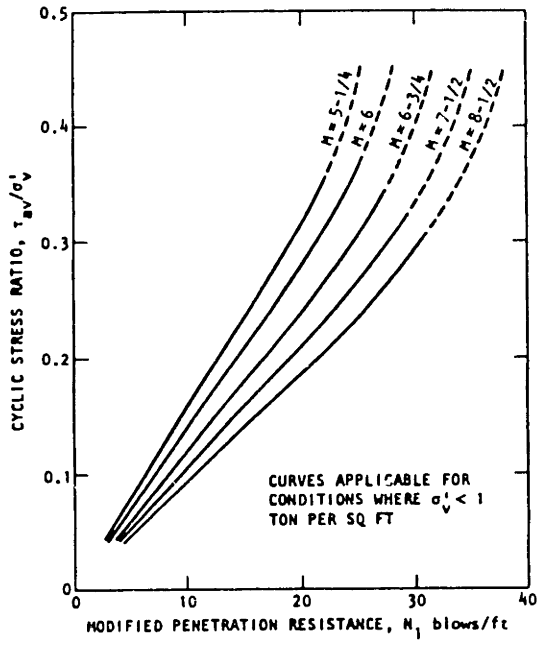
In Chapter 4, a description of the logistic regression method is first presented, and statistical models of liquefaction occurrence are then formulated. Problems of consistency with physical notions of soil liquefaction behavior are considered.

Chapter 5 documents the development of various nonparametric estimators of liquefaction probability. The nonparametric methods were used concurrently with logistic regression primarily to validate the mathematical (parametric) forms assumed in the logit models.

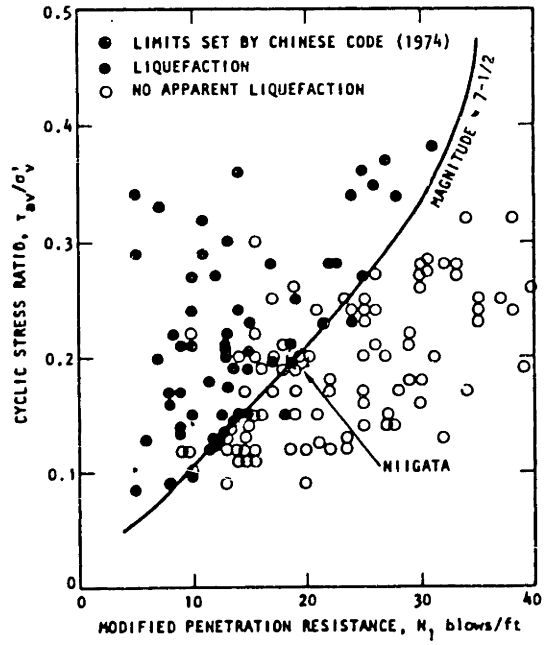
Chapter 6 considers various modifications of the basic logistic regression model, to account for statistical measurement errors and biases in the data. The logistic model is reformulated so that different weights can be assigned to different case study data, adding flexibility to the regression procedure.

In Chapter 7, the statistical models developed in Chapter 4 are considered for implementation in seismic risk analyses. Calculations are presented which illustrate the differences in estimated risk that result for various models developed in this study.

A summary of the thesis, the conclusions drawn from the research, and suggestions for future work, are presented in Chapter 8.



(a) Semi-empirical chart for evaluating liquefaction potential for various earthquake magnitudes (M).



(b) Example data base used for establishing curves in Fig. 1.1(a).

Fig. 1.1 Examples of results from deterministic methods of analyzing liquefaction data. After Seed et al. (1983).

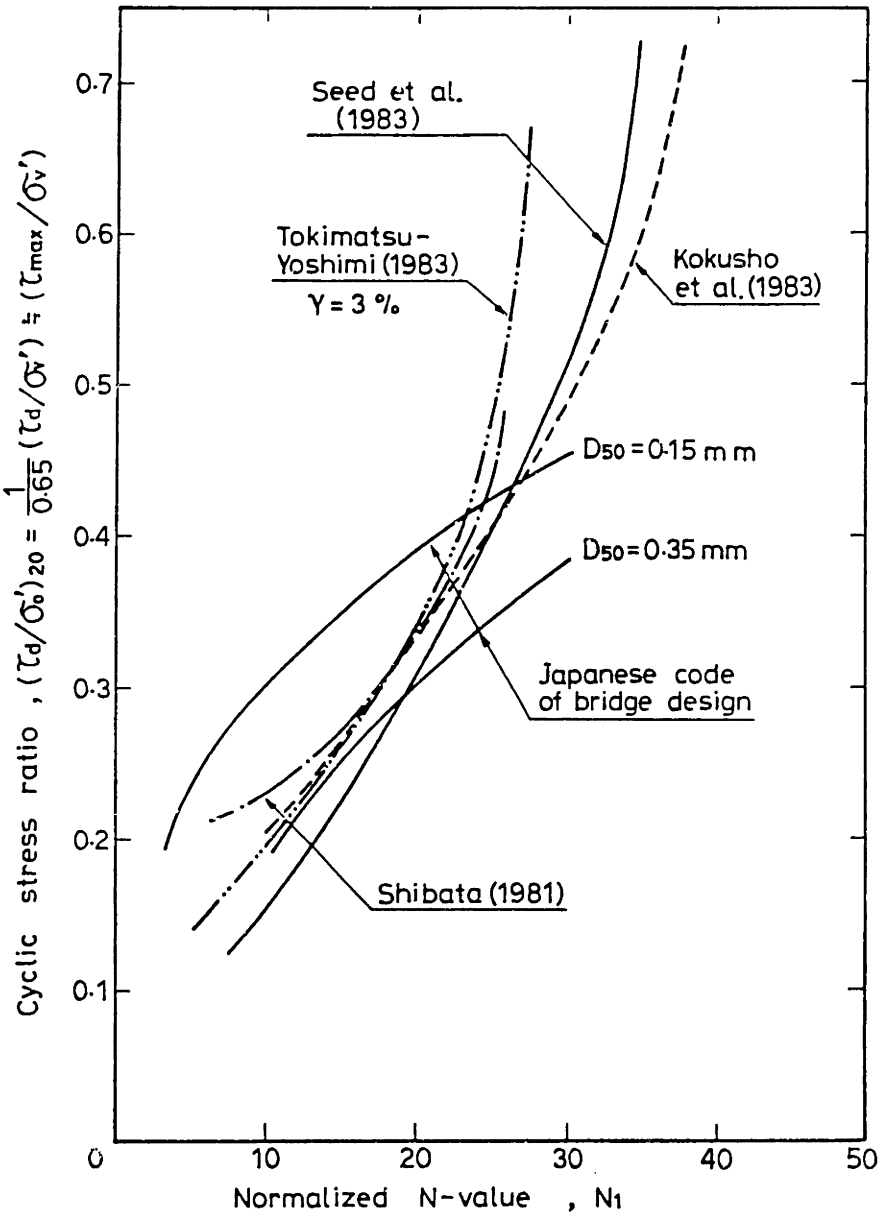


Fig. 1.2 Comparison of various proposed lines separating liquefaction and non-liquefaction behavior in a plot of peak cyclic stress ratio versus N_1 value. From Ishihara (1985).

CHAPTER 2

REVIEW OF PREVIOUS RESEARCH

2.1 Introduction

Liquefaction of soils is a complex and multifaceted problem on which a great deal of research has been done. The purpose of this chapter is to review a relatively small subset of that body of work which specifically pertains to the use of probabilistic and statistical methods in liquefaction analysis. The review is presented within the framework of seismic risk analysis, since the application of most probabilistic or statistical methods are ultimately intended for use within such a framework. For more general reviews, reference should be made to the papers by Yoshimi et al. (1977), Seed (1979) and Ishihara (1985). However, probably the most comprehensive and current review is the report recently published by the National Research Council (NRC, 1985), which resulted from a workshop on liquefaction held at M.I.T. in March, 1985.

Section 2.2 presents an outline of the framework that is commonly used in liquefaction risk assessment and Section 2.3 briefly describes the uncertainties in the determination of earthquake occurrence and intensity. These two sections are meant to provide the background material to put the objectives of this study in perspective. Section 3.3 focuses on the calculation of conditional probability of liquefaction, given the earthquake intensity and soil resistance parameters. In subsequent chapters, where the meaning within the context is clear, the term "conditional" is sometimes omitted.

2.2 The Risk Analysis Framework

The possibility that liquefaction will occur at a given location or project site is a risk that should be taken into account in project planning and design. Probability and statistics provide a systematic way of identifying and quantifying those risks which arise from the following sources:

- Uncertainty on the magnitude and location of earthquakes that can potentially affect the site.
- Uncertainty on the acceleration and duration of ground motion motion at the site, resulting from an earthquake of given magnitude and location.
- Uncertainty on the physical model of soil liquefaction (model uncertainty).
- Uncertainty on the parameters that define such a physical model.

The general steps in earthquake risk analysis, as applied to the problem of liquefaction, are indicated in Fig. 2.1. There are two essential parts. One part deals with the probability that earthquakes of various sizes will occur near the site of interest; this portion is primarily the province of the seismologist. The second part deals with the probability that liquefaction occurs, given certain earthquake characteristics and soil conditions; this is primarily the concern of the geotechnical engineer. The probabilities from these two parts must be combined and summed or integrated over all possible earthquakes to give the probability of liquefaction.

To mathematically formalize the risk analysis procedure, the following generalized notation is introduced:

- Let $\underline{\Psi}$ designate a set or "vector" of earthquake load parameters (e.g. $\underline{\Psi}$ = (a,M), where 'a' is the peak site acceleration and M is the earthquake magnitude).
- Let $\underline{\Omega}$ designate a set or "vector" of liquefaction resistance parameters (e.g. $\underline{\Omega} = ((N_1)_{60}, FC)$, where $(N_1)_{60}$ is the normalized/corrected SPT resistance and FC is the fines content in percent.
- Let Y be an indicator of liquefaction such that Y=1 if liquefaction occurs and Y=0 if it does not occur.

Assume furthermore that the occurrence of earthquakes can be modelled as a stationary Poisson process (See Section 2.3.2). Then the probability that liquefaction occurs at a specific site within a given time period can be expressed as:

$$P(Y=1) = \int_{\underline{\Omega}} \{1 - \exp[-\lambda T \int_{\underline{\Psi}} P(Y=1|\underline{\Omega},\underline{\Psi}) g(\underline{\Psi}) d\underline{\Psi}]\} g(\underline{\Omega}) d\underline{\Omega} \qquad (2.1a)$$

For the usual case of small values resulting from the product λT times the integral over $\underline{\Psi}$, P is well approximated by:

$$P(Y=1) \approx \lambda T \int_{\underline{\Omega}} \int_{\underline{\Psi}} P(Y=1|\underline{\Omega},\underline{\Psi}) \cdot f(\underline{\Omega}) \cdot g(\underline{\Psi}) \, d\underline{\Psi} d\underline{\Omega} \qquad (2.1b)$$

In the equation above:

- $P(Y=1|\underline{\Omega},\underline{\Psi})$ is the conditional probability of liquefaction given $\underline{\Omega}$ and $\underline{\Psi}$.
- $g(\underline{\Psi})$ is the joint probability density function of the earthquake load parameters.
- $f(\underline{\Omega})$ is the joint probability density function of the soil resistance parameters.

- λ is the overall rate of earthquake occurrence (e.g. earthquakes per year) from all seismic sources within a given neighborhood of the project site.
- T is the time period under consideration (e.g. project life in years).

Eqn. 2.1 is derived based on modelling liquefaction as a "Poisson process with random selection," i.e. liquefaction does not occur with every generic earthquake, but only for randomly "selected" earthquakes. The rate of liquefaction occurrence is thus a fraction of the earthquake occurrence rate λ. This fraction corresponds to the probability that a generic earthquake will result in liquefaction and is represented by integral over $\underline{\Psi}$ in Eqn. 2.1(a). Details of the derivation may be found in McGuire et al. (1978,1979). Eqn. 2.1 can be easily adapted to account for a multiplicity of source zones with different occurrence rates λ and densities $g(\underline{\Psi})$; see Atkinson et al. (1984). The integrals in Eqn. 2.1 are usually evaluated numerically and represent the generalized summation of all the probabilities associated with each of the possible combinations of earthquake magnitudes, locations, and site conditions. As a practical matter, the integral over $\underline{\Omega}$ is usually not calculated, due in part to the difficulty of quantifying $f(\underline{\Omega})$.

The above general formulation for assessing the probability of liquefaction at a site is adapted from the seismic hazard methodology originally proposed by Cornell (1968). Methods of liquefaction risk analysis have been proposed by McGuire et al. (1978, 1979), Yegian and Whitman (1978), Haldar and Tang (1978), Yegian and Vitelli (1981a, 1981b), Chameau and Clough (1983), Atkinson et al. (1984), and Kavazanjian et al. (1985). All of these methods have the above framework

in common, but differ in the evaluation of the functions $P(Y=1|\underline{\Omega},\underline{\Psi})$, $f(\underline{\Omega})$ or $g(\underline{\Psi})$ or in the integration algorithm.

It should be noted that the probability of liquefaction is always evaluated in the context of a specified time interval. In a seismically active region, the probability of liquefaction at a susceptible site increases with the time of exposure (T) to the earthquake risk. Though probabilities of liquefaction may in some instances be negligibly small, zero risk is only theoretically possible for $T = 0$.

2.3 Earthquake Load Parameters

The variability in the earthquake load parameters is generally dealt with in the context of seismology and seismic hazard models, such as those by Cornell (1968) and Der Kiureghian and Ang (1977). The quantifications of earthquake hazard is a complex task and is described briefly here in terms of how it interfaces with various aspects of liquefaction risk analysis.

2.3.1 Seismic source zones

The first step in evaluating seismic hazard is in the definition of regions where earthquakes originate. Seismic source zones can be faults or regions that can be mapped based on past seismicity. It is usually assumed that earthquakes of any given magnitude have an equal chance of occurring anywhere within each seismic zone. In the absence of well defined earthquake mechanisms in many geographic regions or for the sake of simplicity of analysis, it is sometimes assumed that the seismic source zone is circular and centered at the project site, as shown in Fig. 2.2. However, such simplistic procedures are becoming less common.

In more sophisticated models or algorithms, where seismicity is relatively well defined, several irregularly shaped seismic source zones may be defined as shown in Fig. 2.3.

2.3.2 Rates of Seismic Activity

The rate of seismic activity in a source zone is usually determined based on the analysis of historical seismicity. It is customary to assume that the occurrence of earthquakes is a stationary Poisson process in time. "Stationary" implies that the rate of earthquake occurrence (expected number of earthquakes per year) is constant over time. "Poisson" implies that the occurrence of future earthquakes is independent of earthquakes that have occurred in the past. These two assumptions are generally not strictly valid. In the framework of plate tectonics, after the release of strain in the earth's crust during a large earthquake, a certain period of time is required to build up enough strain to again cause another large earthquake rupture. Though the assumption of a stationary Poisson model for earthquake occurrence is frequently made in practice, recent advances in seismology and earthquake prediction may offer better representations of earthquake occurrences (Patwardan and Kulkarmi, 1980; Kiremedjian and Anagnos, 1984; Agnanos and Kiremedjian, 1984; Schwartz and Coppersmith, 1984; Sykes and Nishenko, 1984).

2.3.3 Seismic Source Mechanisms

There are two major seismic risk models in use, which differ in the modelling of the nature of the seismic source. The Cornell (1968) model assumes that the energy released during any earthquake is initially

concentrated at a point and propagates outward from that point. In contrast, the model by Der Kiureghian and Ang (1977) assumes linear sources (faults) and that the energy released is not concentric about a point, but rather propagates outward from a finite length of fault rupture. Thus, the largest contribution to earthquake intensity at a site may not be related to the location of the initial fault rupture, but rather to the slip that occurred closest to that site. The Der Kiureghian and Ang (1977) model is applicable to seismic regions such as California, where the strike-slip faults causing earthquakes are well defined and have been mapped extensively. By contrast, in regions where the sources and perhaps even the nature of seismic activity is not well defined, the Cornell (1968) model is probably more appropriate.

2.3.4 Local versus Global Intensity Measures

Most liquefaction risk analyses describe the intensity of earthquake shaking at the site in terms of peak acceleration and duration. This characterization (sometimes called the A&D approach) is represented by the methods proposed by McGuire et al. (1978, 1979), Haldar and Tang (1979), Chameau and Clough (1983), Atkinson et al. (1984) and Kavazanjian et al. (1985). Acceleration is usually obtained from an attentuation law, and duration (usually stated in terms of number of equivalent load cycles) is given as a function of magnitude. However, there is uncertainty in the relation between duration and magnitude and in the attenuation of acceleration with distance. Formal treatment of uncertainty in attentuation was introduced by Cornell (1971) and is commonly incorporated into earthquake hazard analyses. For example, Atkinson et al. (1984) consider a log-normal dispersion of acceleration

around the median value from an attenuation relationship.

An alternative approach, using "global" parameters such as magnitude (M) and distance (R), sometimes called the M&R approach, is represented by the methods of Yegian and Whitman (1978), Yegian and Vitelli (1981a, 1981b) and Youd and Perkins (1978). The M&R methods are more direct and arguably easier to use within a risk analysis framework, especially for applications such as large-scale mapping of liquefaction potential. Specifically, these methods eliminate the need for the intermediate step of applying an attenuation law with its attendant considerations of uncertainty. Furthermore, the appropriate choice of an attenuation law from among the many available requires considerable judgement. However, use of M&R methods requires that the probability of liquefaction be related directly to magnitude and distance, which tends to obscure local site effects that may influence the occurrence of liquefaction. This will be more clearly shown by the results presented in subsequent chapters.

2.4 Conditional Liquefaction Probability

2.4.1. Types of Formulations

The conditional liquefaction probability $P(Y=1|\underline{\Omega},\underline{\Psi})$ is the probability that a site liquefies, given a specified intensity of earthquake shaking and soil deposit parameters at the site. Some liquefaction risk procedures, such as those presented by McGuire et al. (1978, 1979), Youd and Perkins (1978) Davis and Berrill (1982), and Atkinson et al. (1974) do not account for any uncertainty in the prediction of liquefaction if the load and resistance vectors $\underline{\Omega}$ and $\underline{\Psi}$ are

given. That is, the conditional probability function $P(Y=1|\underline{\Omega},\underline{\Psi})$ is assumed to take on values of one or zero (liquefaction or no liquefaction), which represents a purely deterministic formulation. All the uncertainty in their liquefaction risk assessments arises from uncertainty of future earthquake occurrence and on the local soil properties.

Actually there is uncertainty in any method used to determine whether or not liquefaction will occur given $\underline{\Psi}$ and $\underline{\Omega}$. All analyses, whether they are represented by a simple graph or embodied in the form of a complex computer program, involve some theoretical model for the physical behavior of a site. Model uncertainty arises as a result of simplifications or assumptions made in formulating the model, and even if all the input parameters were known precisely, there would still be some uncertainty about the model prediction of liquefaction behavior. An example is the original semi-empirical liquefaction model proposed by Seed, Arango,and Chan (1975) which is represented by a plot of the cyclic stress ratio CSR (τ/σ_o') versus the corrected SPT resistance N_1 (i.e. the same format as Fig. 1.1). In that model, no provisions were made for duration effects of earthquakes, the differences in methods of obtaining N_1, or the effects of fines content. The neglect of these factors was the source of uncertainties in the the original model. Even though these factors have since been incorporated into later versions of the model, some model errors still remain. Examples of possibly significant variables not taken into account include the frequency content of strong motion shaking, and specific drainage conditions in a soil profile. In addition, there are uncertainties in the inputs to any model. For

example, it is well known that the measurement of N_1 is inexact and is subject to considerable statistical error.

Approaches which have been used to obtain the conditional probability of liquefaction in other than yes/no terms can be categorized as probabilistic or statistical:

- Probabilistic methods involve the use of behavioral models for liquefaction at the site, estimating the inherent uncertainties in the parameters of the model, and then propagating these uncertainties through the model.
- Statistical methods generally deal with the extraction of information from data; in the case of the liquefaction problem, "data" usually refers to field observations of instances of liquefaction or non-liquefaction.

The following sections describe in more detail the application of deterministic rules, probabilistic methods, and statistical methods that have been used to calculate the conditional probability of liquefaction. It should be noted that deterministic rules for predicting liquefaction are sometimes derived using statistical methods. In these cases, statistics is used merely as a data analysis tool and not as a means of constructing non-deterministic models.

2.4.2 Deterministic Rules

An example of the use of a deterministic rule is in the risk analysis procedure of Atkinson, Finn, and Charlwood (1984). Basically, these authors use a version of the deterministic liquefaction model originally formulated by Seed and Idriss (1971) and later modified in various ways, e.g. see Seed, Idriss, and Arango (1983). In the procedure

of Atkinson et al (1984), the deterministic rule can be written as follows:

$$P(Y=1 \mid CSR, M, N_1) = \begin{cases} 1 \text{ if } CSR > CSR_{crit} \\ 0 \text{ otherwise} \end{cases} \qquad (2.2)$$

in which CSR is the cyclic stress ratio (also frequently denoted as $\tau/\bar{\sigma}_v$ or τ/σ_o'), M is the earthquake magnitude, and N_1 is the corrected SPT resistance. The cyclic stress ratio CSR is calculated by the formula

$$CSR = 0.65 \frac{a}{g} \frac{\sigma_v}{\bar{\sigma}_v} r_d \qquad (2.3)$$

where a is the peak ground acceleration, g is the gravitational acceleration, σ_v and $\bar{\sigma}_v$ are the total and effective overburden stresses, and r_d is a depth reduction factor. [See Chapter 3, Sec. 3.3.2 for more details regarding the Seed-Idriss method.] The quantity CSR_{crit} is a function of M and N_1 and is defined by boundary lines which, for any given M, separate liquefaction and non-liquefaction events in the CSR versus N_1 plane; see Fig. 2.4.

The Seed-Idriss analysis, which involves the calculation of CSR, is an example of an A&D liquefaction model. The cyclic stress ratio is directly related to acceleration, and although the magnitude M is involved in the analysis, it actually enters as a measure of earthquake duration. An example of a deterministic rule involving M&R model in a liquefaction risk procedure is the method proposed by Youd and Perkins (1978). In their formulation, the conditional liquefaction probability is:

$$P(Y=1|M,R_{EP}) = \begin{cases} 1 \text{ if } R_{EP} \leq R_{crit} \\ 0 \text{ otherwise} \end{cases} \quad (2.4)$$

R_{EP} is the epicentral distance and enters in this simplistic formulation as an equivalent resistance parameter. Note that this model does not take into account the soil resistance at the site of interest. R_{crit} is a function of magnitude and corresponds to a boundary line separating liquefaction and non-liquefaction in a plot of M versus R_{EP}, as shown in Fig. 2.6.

2.4.3 Probabilistic Models

Haldar and Tang (1979) have used a first order second moment (FOSM) method applied to the Seed and Idriss (1971) procedure to obtain the conditional probability of liquefaction. Basically, this involved estimating the uncertainty on the parameters of the Seed-Idriss model (i.e. a, σ_v, $\bar{\sigma}_v$ and r_d in Eqn. 2.3) and propagating these uncertainties through the model. A more sophisticated FOSM model has been presented by Fardis and Veneziano (1982) incorporating the effects of pore pressure diffusion, soil stiffness reduction, and variation of soil properties within a stratum. In both models, the assumption of normality or log-normality of load and resistance parameters is used in estimating the conditional probability of liquefaction.

Probabilistic analyses based on pore pressure generation models have been presented by Chameau and Clough (1983). The accumulation of pore pressure is calculated using a nonlinear formulation, based either on laboratory data or on a basic constitutive physical model. They find the conditional probability $P(Y=1|\underline{\Omega},\underline{\Psi})$ as the probability that the pore

pressure ratio r_u equals 1, which in turn is calculated assuming random arrivals of shear stress (or equivalent acceleration) peaks between positive zero crossings of the earthquake record. Chameau and Clough (1983) suggest that the distribution of the shear stress peaks can be modelled as beta, gamma, Rayleigh or exponentially distributed. The number of positive zero crossings is related to duration and spectral content and may also be considered random, as is done in the application of Chameau and Clough's method by Kavazajian et al (1985).

Earlier, Donovan (1971) proposed a similar model, assuming the variation of peak amplitudes of earthquake shaking to be distributed as a Rayleigh probability function. In his method, the effects (damage) due to the various peak accelerations in an earthquake record are calculated based on a total stress model of cyclic loading behavior, and are summed or integrated with a weighting function proportional to the frequency of occurrence of the various peaks (Miner's Linear Damage Criteria). However, the method produces a deterministic criteria for liquefaction and not a conditional liquefaction probability.

There are two major drawbacks to probabilistic methods. First, as indicated by the various examples, there is a fairly diverse range of models that have been used in application, and there is no easy way to compare the actual performance of these models. Rather, comparisons have to be made in terms of the features incorporated in each of the models. Secondly, the application of probabilistic methods necessitates the assumption and specification of the probability distributions of the inputs to the model. Whether they are second moment parameters (means and standard deviations) or the entire probability distribution, the

specification is not a trivial task and requires some judgement.

2.4.4. Statistical Analyses

The predominant problem that has been addressed in the context of statistical methods is that of how to draw the "best" boundary separating liquefaction and non-liquefaction behavior, e.g. on a diagram such as Fig. 2.6. This is known in statistics as a problem of classification or discrimination. This problem has been treated by Christian and Swiger (1975) using empirical data on site liquefaction behavior and a statistical method known as linear discriminant analysis (e.g., see Johnson and Wichern, 1982). Applications of discriminant analysis for studying liquefaction have also been reported by Tanimoto and Noda (1976), Tanimoto (1977), Xie (1979), Wang et al. (1980), Davis and Berrill (1981, 1982, 1985) and Gu and Wang (1984).

Discriminant analysis is based on the assumption that there are distinct statistical populations, i.e. of liquefied and non-liquefied sites, from which random samples have been obtained. From these samples, a criterion is developed to classify future observations as belonging to one of those populations. The rule for classification derived by Christian and Swiger (1975) can be stated as:

- Classify as liquefaction if $D_r \leq KA^{0.31069}$
- Classify as non-liquefaction if $D_r > KA^{0.31069}$

where D_r is the relative density of the soil and A is a modified site acceleration. This rule can be plotted as a series of discriminant lines in the A versus D_r plane, as shown in Fig. 2.6. The factor K is a function which depends on the probability of misclassifying a data point

as non-liquefaction when in actuality it is a liquefaction point. This probability is denoted as P in Fig. 2.6, and expresses the confidence level or degree of conservatism with which one can make design decisions based on the associated dividing line. The probability P in Fig. 2.6 is not the conditional liquefaction probability $P(Y=1|\underline{\Omega},\underline{\Psi})$. The result of discriminant analysis is essentially the same as a deterministic rule (see Section 2.4.2). To apply the Christian and Swiger (1975) result in the context of evaluating liquefaction probability using Eqn. 2.1, one would have to interpret the rule as:

$$P(Y=1|D_r,A) = \begin{cases} 1 \text{ if } D_r \leq KA^{0.31069} \\ 0 \text{ if } D_r > KA^{0.31069} \end{cases} \quad (2.5)$$

A problem with employing discriminant procedures in the analysis of liquefaction is that the data do not represent random samples from liquefied and non-liquefied sites; see Easterling and Heller (1976). Underlying assumptions of normality of the data are a feature of most discriminant analysis methods, and these assumptions are also obviously not satisfied. Nonparametric methods of discriminant analysis have been developed (Hand, 1981, 1982) which do not require a normality assumption, but these methods have not been applied to the liquefaction problem. Furthermore, it is meaningless to a certain extent to talk of "populations" of liquefied sites and non-liquefied sites with well-defined distributions of explanatory variables. For example, the probability distributions of acceleration and N_1 for liquefied sites will typically be different in different regions of the world, and the meaning

of "population" is thus not clear.

The results of discriminant analysis are also highly sensitive to the distribution of the data. A one-dimensional schematic to illustrate this is shown in Fig. 2.7. Suppose one has an "original" data set from which a discriminant criterion is determined as shown in Fig. 2.7(a) by the intersection of the tails of normal distributions fitted to the two sets of liquefaction and non-liquefaction data. [This method of obtaining a discriminant criterion is not used in practice, but is presented as a simple illustration.] Then, suppose that additional information on liquefied sites is acquired due to a series of high intensity earthquakes as shown in Fig. 2.7(b). With this additional data, a new discriminant criterion would be obtained, even though no changes in the data base have occurred in the central region of the plot in Fig. 2.7. This illustrates the sensitivity of the discriminant criterion to extremal data, which contribute essentially little "information" near the critical region in the analysis.

Yegian and Whitman (1978) presented a different classification method (developed by Yegian, 1976) which they termed the "least squares of the misclassified points". In essence, this method finds the line which best separates liquefaction from non-liquefaction points based on minimizing the sum of the squared distances between the misclassified points and the boundary line. Note that an advantage of this procedure over discriminant analysis is that it is not sensitive to extremal data points. Atkinson et al. (1984) have used this method in part of their analyses, although their algorithm also accommodates the use of the more conservative lower bound curves from Seed (1979).

The problem with classification methods is that the discriminant criteria still gives a deterministic yes/no type answer to whether liquefaction will or will not occur at a site rather than a continuous conditional probability $P(Y=1|\underline{\Psi},\underline{\Omega})$ that varies smoothly between 0 and 1. Yegian (1976) and Yegian and Vitelli (1981a,1981b) obtain a conditional liquefaction probability by assuming the inverse of the factor of safety (denoted LPI) obtained using the classification criterion to be log-normally distributed. The mean and standard deviation of ln(LPI) are then calculated from an analysis of the errors of the inputs, treating LPI essentially as a deterministic criterion.

2.5 Commentary

This study is primarily concerned with the estimation of the conditional probability of liquefaction $P(Y=1|\underline{\Omega},\underline{\Psi})$. Methods are used which avoid many of the problems of probabalistic and statistical approaches that have been previously employed. Instead of treating the statistical analysis of liquefaction data as a problem of classification, the estimate of $P(Y=1|\underline{\Omega},\underline{\Psi})$ is obtained through regression using an indicator Y of liquefaction as a binary (zero/one) response variable. A more detailed comparison of classification versus binary regression is presented in Chapter 4.

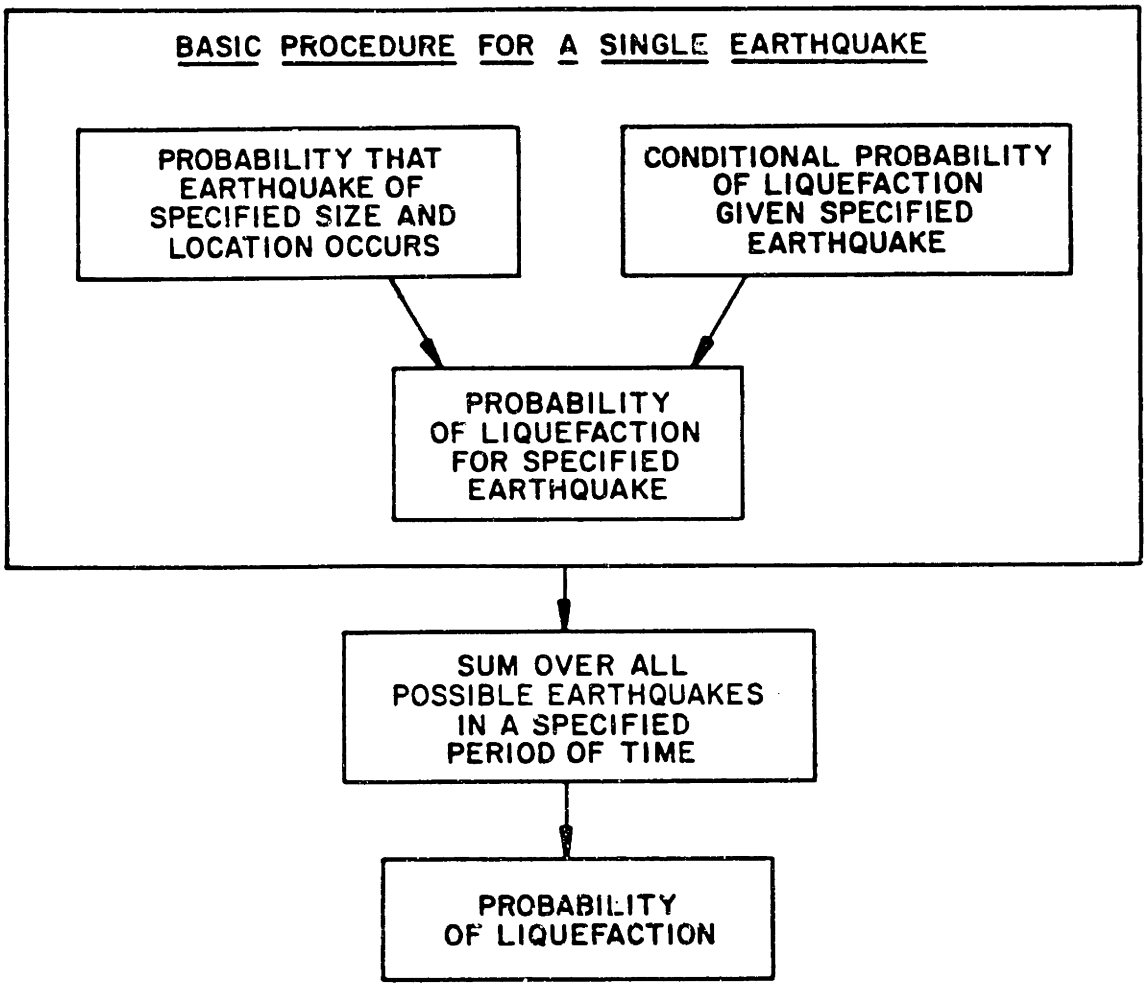


Fig. 2.1 Schematic of the steps in a probabilistic liquefaction risk analysis.

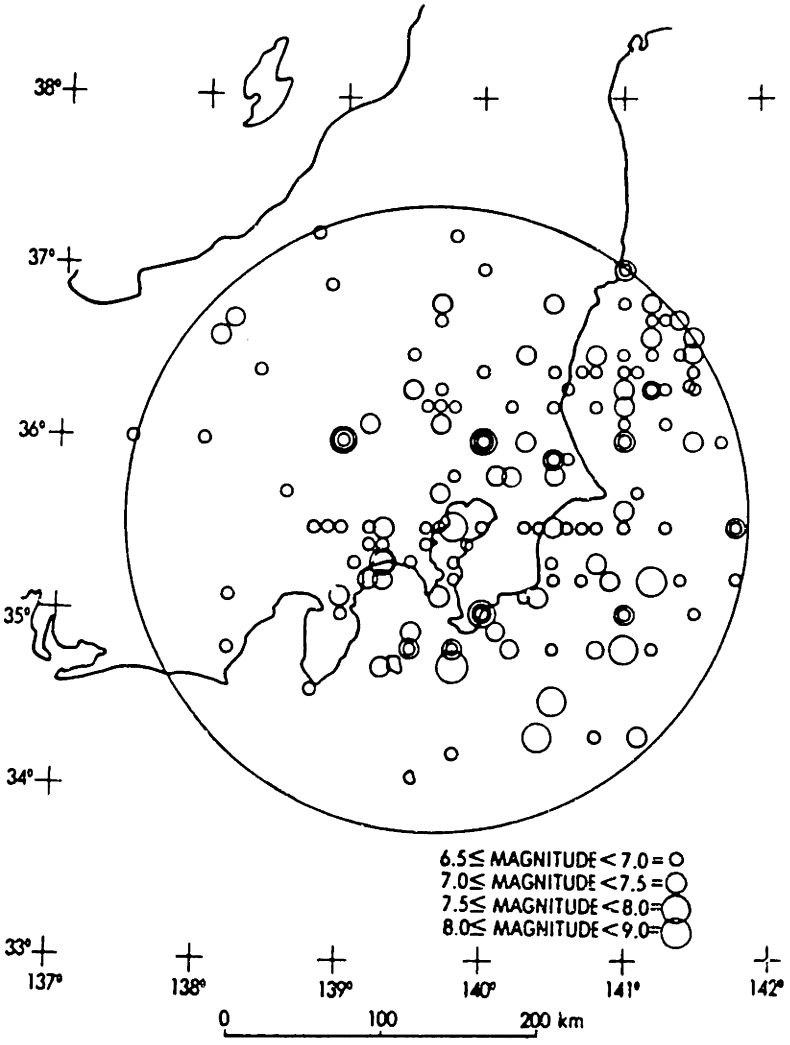


Fig. 2.2 Circular seismic source zone centered at project site for liquefaction risk analysis. From McGuire et al. (1978).

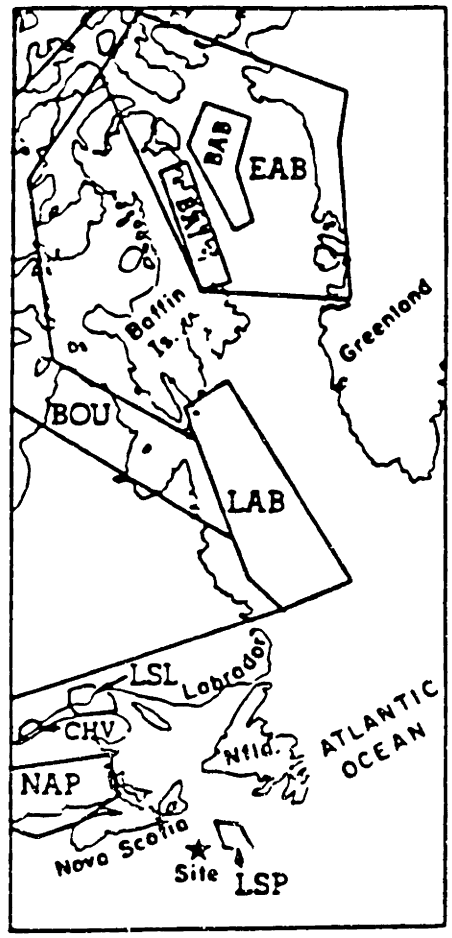


Fig. 2.3 Complex set of irregularly shaped seismic zones (labeled LAB, BOU, etc.) for liquefaction risk analysis. From Atkinson, Finn and Charlwood (1984).

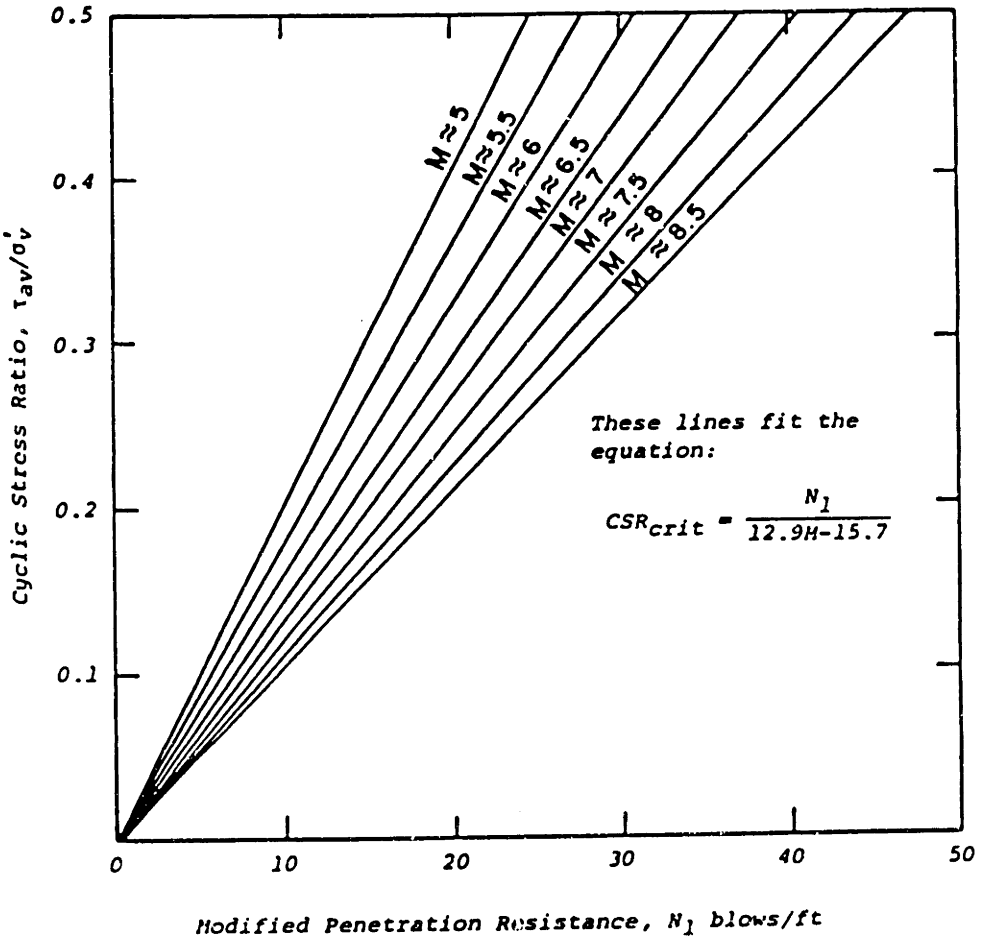


Fig. 2.4 Deterministic criteria for liquefaction used in the risk analysis procedure of Atkinson, Finn, and Charlwood (1984).

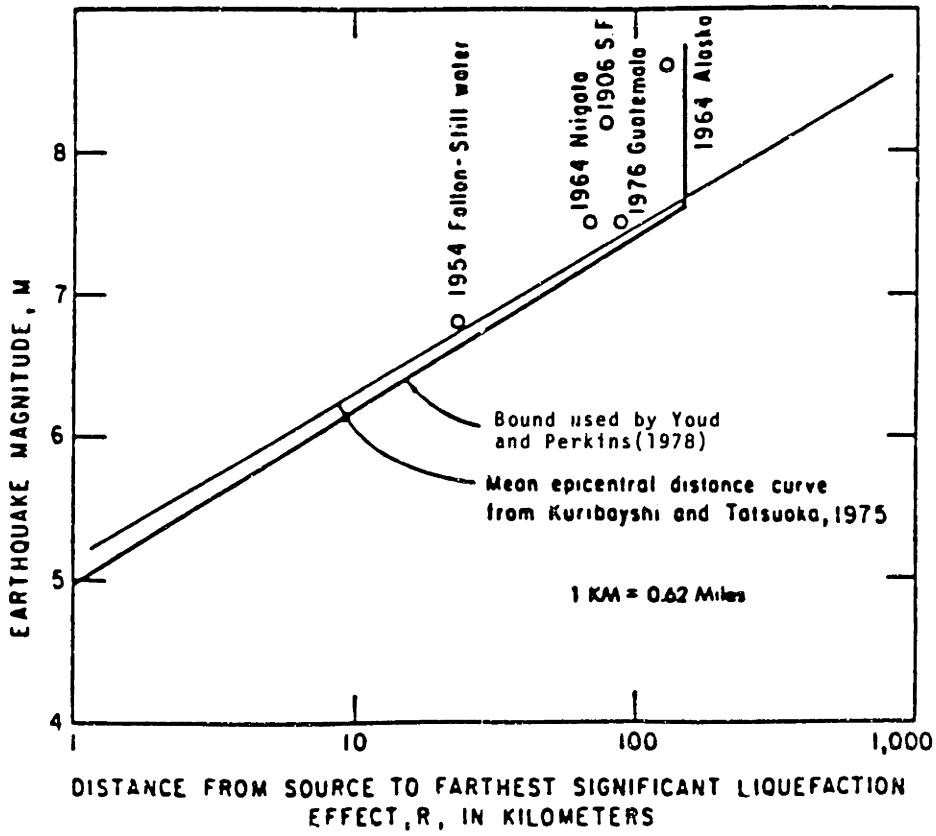


Fig. 2.5 Deterministic criteria for liquefaction proposed by Youd and Perkins (1978) for mapping liquefaction risk.

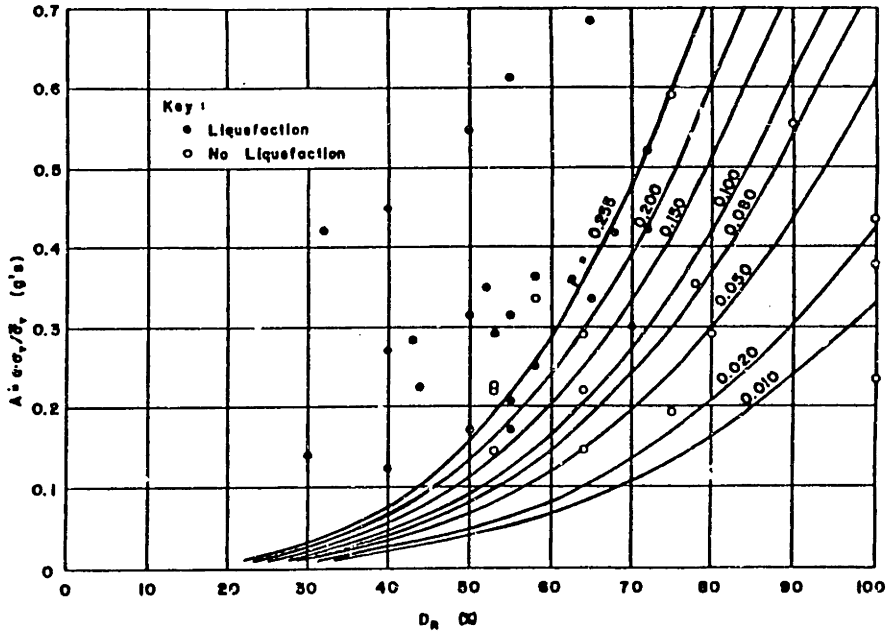


Fig. 2.6 Example result from discriminant analysis of liquefaction data. From Christian and Swiger (1975).

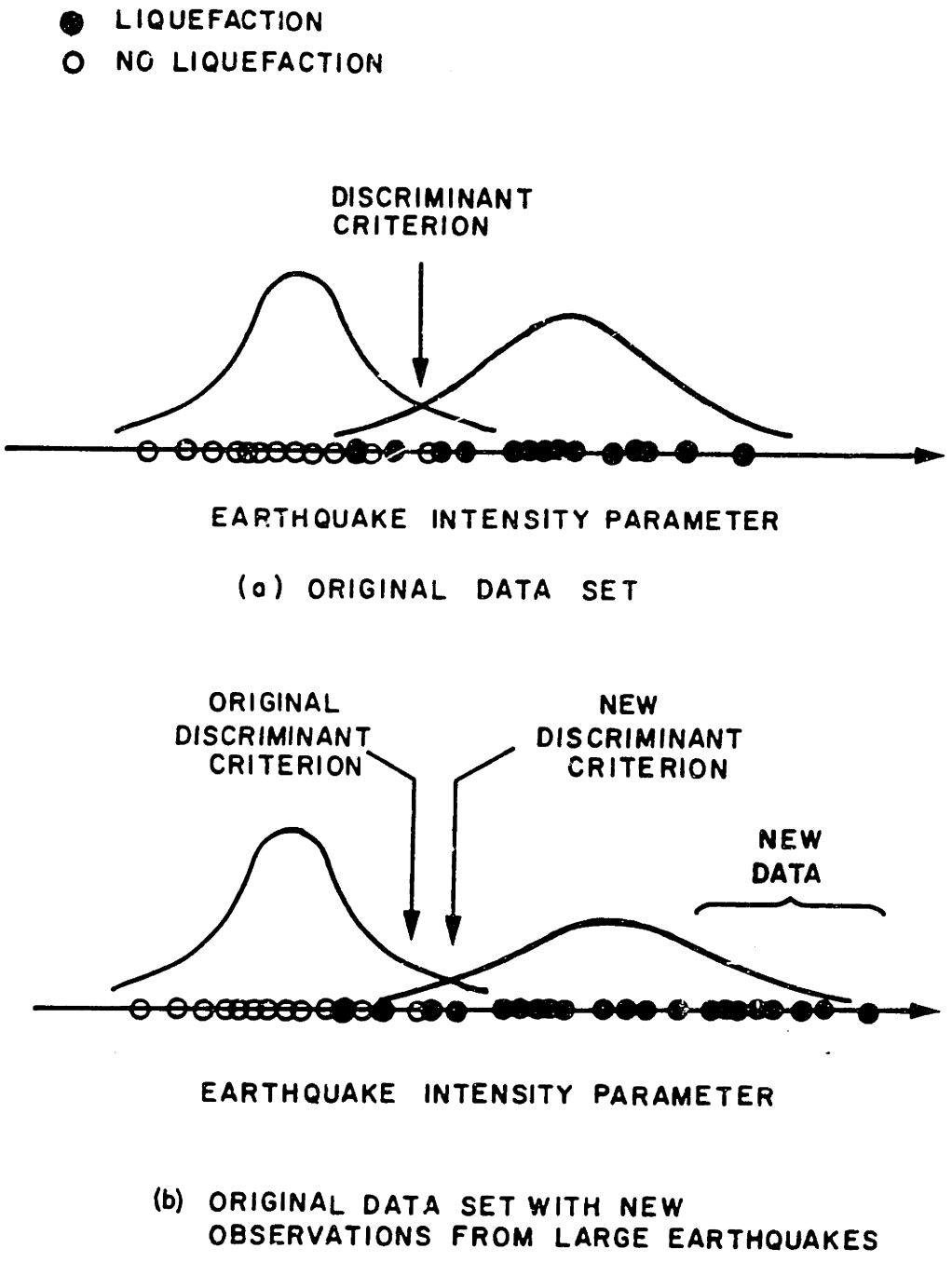


Fig. 2.7 Schematic of sensitivity of discriminant analysis to changes in data base.

CHAPTER 3

LIQUEFACTION DATA CATALOG

3.1 Introduction

This chapter describes the main features of the data base compiled for the statistical analysis of field liquefaction behavior. The data catalog is presented in Appendix A and consists of a total of 278 case studies, with 114 cases representing sites that liquefied during earthquakes and 164 cases representing sites which did not. The catalog was compiled from the synthesis of eight previously published "source catalogs" of similar format, and enhanced by individually re-examining many of the case studies. Several recently published case studies were also incorporated.

The synthesized catalog contains data that were available to the writer as of the end of October, 1984, which represents the closing date of this compilation. Attempts were made to obtain additional data by various means, including correspondence with other researchers, but many did not or could not respond by the closing date.

The format of the compiled catalog has been influenced by its intended use in logistic regression, which will be described in Chapter 4. For example, one feature of the catalog is the utilization of binary variables, i.e. variables that take on values of 0 or 1 to indicate different categories of data. An example is a variable which equals 0 if the soil is a clean sand and equals 1 if the sand is silty or contains a significant amount of fines.

It is realized that no tabulation of liquefaction case studies can be completely exhaustive or free of error. Also, there are undoubtedly biases introduced by any compiler of such catalogs, and future researchers may wish to re-examine the case studies and to correct or eliminate certain data. Thus, the catalog is formatted so that it may be easily revised or expanded to accommodate new data. Commentaries are provided in Appendix A for selected case studies, in order to illuminate the uncertainties and judgements that had to be made for particular cases during the compilation of the data. A more general discussion of the problems that have been encountered during compilation is presented here.

3.2 Previous Catalogs

In 1971, Seed and Idriss published a set of data on 35 sites that did or did not liquefy during earthquakes, including for each site, the values of a few soil parameters, earthquake magnitude and ground motion characteristics. In the same year, Whitman (1971) published a smaller catalog of 13 case studies, some of which were an independent evaluation of the Seed and Idriss cases. A more comprehensive catalog was subsequently published by Seed, Arango, and Chan (1975), which expanded the Seed and Idriss (1971) catalog to 38 cases, and included more parameters for each case study. Additional catalogs have been compiled by Yegian (1976), Yegian and Vitelli (1981a), Xie (1979), Davis and Berrill (1981), Tokimatsu and Yoshimi (1983), and Seed et al. (1984).

Table 3.1 summarizes the "source catalogs" that were used for the synthesis of the data base used in this study. Other catalogs exist in the literature, but they are either of a format that was not useful for

the present compilation, or their data are a subset of one of the "source catalogs". From Table 3.1, it can be seen that the characteristics of previously published catalogs have considerable variation. The Seed, Idriss, and Arango (1975) data base, while comparatively limited in the total number of case studies, provides the most comprehensive description of earthquake load parameters. The somewhat larger catalog of Davis and Berrill (1981) provides only sparse details on earthquake and site soil parameters. The Tokimatsu and Yoshimi (1983) catalog and, to a lesser extent, the Yegian and Vitelli (1981) catalog provide the largest amount of soil data.

A major difference in the philosophy of earthquake characterization may be seen in comparing the catalogs of Yegian (1976), Yegian and Vitelli (1981a), and Davis and Berrill (1981) with other data sets. This difference is in the choice of global earthquake-intensity measures (e.g. magnitude M and distance R) versus local measures (e.g. peak site acceleration and duration of shaking). Both types of data are used in this study.

The data catalogs of Table 3.1 contain biases and are imperfect to various extents. The main sources of deficiency are:

- Non-uniform Quality of the Data. Some of the data reported in the catalogs are from earthquakes that occurred more than a century ago. Clearly, these older data should be regarded as less reliable than information from recent events. Another cause of potentially poor quality data is the lack of earthquake recording networks in some geographic regions.
- Lack of Statistical Independence Between Case Studies. In the

Yegian (1976), Yegian and Vitelli (1981) and Tokimatsu and Yoshimi (1983) catalogs, there are instances of two or more case studies obtained using data at different depths from the same boring. Also, many cases are obtained from several borings at the same site. The physical proximity of the data raises the question of independence between the cases. It is also a common practice to use the same boring data as a series of successive case studies in different earthquakes. For example, Seed et al. (1975) use the same soil data in different case studies for the 1802, 1877 and the 1964 earthquakes in the Niigata region. Some correlation between these case studies is more than likely.

- Non-Proportional Sampling of Liquefaction Versus Non-Liquefaction Sites. In general, liquefaction sites tend to be studied in more detail and are more extensively reported than non-liquefaction sites. Hence, the proportion of liquefied to non-liquefied sites tends to be higher in the catalogs than in reality. This source of bias, from reported relative frequency, affects the estimation of liquefaction probability.
- Measurement Errors. Errors of this type are present, for example, in the estimation of earthquake magnitude, epicentral distance, and the SPT N-value.
- Difficulties of Site Characterization. In most of the liquefaction catalogs, sites are characterized by only a single SPT N-value. How this value is selected to be representative of a boring profile or even of the entire site is not always clear nor consistent.

- Lack of Differentiation Between SPT Data Obtained Before and After Earthquake Occurrence. Although changes in density and SPT resistance are purported to occur as a result of earthquake shaking (e.g. Koizumi, 1966), these changes are often ignored. The implicit assumption usually made is that post-earthquake values of SPT resistance adequately represent the site before the earthquake occurred.
- Lack of Differentiation Between N-values obtained using Various SPT Methods. With the exception of the Tokimatsu and Yoshimi (1983) and Seed et al. (1984), the various source catalogs do not differentiate between the "standard" rope and pulley method and the "free fall" methods of performing the SPT. Tokimatsu and Yoshimi also indicate that a difference exists between the results of the SPT using Japanese and non-Japanese drilling methods. Seed et al. (1984) have provided corrections to account for these factors, and their methodology has been incorporated in the synthesized catalog.

Attempts were made to compensate for some of the above sources of deficiency, though it was not possible to remove all the imperfections. An important feature of the synthesized catalog is that the sources of deficiency are noted and recorded in the form of numerical codes. Analyses of the potential effects of these imperfections on the estimates of liquefaction probability are presented in subsequent chapters. However, the problems of site characterization and statistical independence have affected the compilation of the data to a greater extent, and are discussed in Section 3.4, along with some other problems

specific to data compilation.

3.3 Synthesized Catalog Description

Appendix A of this thesis presents the catalog of liquefaction case studies in six tables:

Table A.1 - Case Identification and Qualitative Attributes

Table A.2 - Edited/Enhanced Data

Table A.3 - Previously Published Catalog Identification Codes

Table A.4 - Data from Previously Published Catalogs

Table A.5 - Case Source Reference Guide

Table A.6 - Magnitude Scales for Tabulated Earthquakes

Also included in Appendix A is a bibliography of "source references", where most of the original data and boring logs for the case studies can be found. Not all of the source references were available to the writer at the time this compilation was completed. Commentaries on selected case studies are presented to indicate specific difficulties, assumptions, or judgements that had to be faced in compiling the data.

Table A.1 identifies each case study with a code that associates it with an earthquake and the particular site of the case study. Whether liquefaction occurred or not is indicated by a binary (zero/one) variable. In addition, there are binary codes that reflect the qualitative aspects of the data presented in Table A.2, including descriptions of:

- Whether the case study is located in one of the three dominant geographic regions (Japan, California, or China) which make up the bulk of the case study data.
- Whether the reported epicentral distance is based on

instrumental data or on the location of maximum reported intensity.

- Whether the local site acceleration was obtained from attentuation relationships/intensity data, or from a strong motion recorder at or near the site.
- Whether the N-value was obtained from the SPT performed using the rope-and-pulley or a free-fall method, from correlations with static cone penetration data or by other means.
- The method used to choose a representative SPT N-value that characterizes the soil profile at the site (see Section 3.3.4).
- Whether the case involved a slope stability or embankment failure problem. (Level ground or level ground areas with building foundations is the default category.)
- Whether artesian conditions may have existed at the site. (There are only 2 cases where this occurred.)
- Whether the boring data were obtained before or after the earthquake.

Table A.1 also includes codes which indicate the case studies that may be correlated with each other by virtue of being located at the same site or being based on the same boring data.

Table A.2, Edited/Enhanced Data, presents the quantitative variables associated with each case study. All of the variables which have been reported in the source catalogs, except for relative density and hypocentral distance, are reported in Table A.2. Relative density is not listed, because in-situ estimates of this quantity are usually based on the interpretation of SPT resistance. The hypocentral distance can

theoretically be calculated given the focal depth of the earthquake and the epicentral distance. Binary codes for soil description based on visual classification are tabulated in Table A.2. These are particularly useful in cases where grain size analyses have not been reported, and allows the discrimination between cases, for example, of sites where the soil is visually classified as a silty sand versus a clean sand. In addition to the binary liquefaction indicator, Table A.2 includes a code for the classification of the severity of liquefaction proposed by Tokimatsu and Yoshimi (1983). These codes are not used in this study but are included for possible use in future research. Also, it should be noted that there are "missing data" codes to indicate data that are not reported for certain case studies, e.g. fines content data is not available for every case.

Tables A.1 and A.2 are the main products of the synthesis and compilation of data from the source catalogs and contain the data analyzed in this study. Tables A.3 and A.4 (which represent an intermediate stage in the compilation of Tables A.1 and A.2) are included for documentation purposes only. Table A.4 presents the data in each case study, as given in the original source catalogs, except that in some cases, the modified SPT resistance N_1 has been changed to reflect standardization using the overburden correction factor proposed by Liao and Whitman (1986) (see Sec. 3.4.3). If the data in more than one source catalog are the same, only one set of case data is reported in Table A.4 to avoid repetition. For each case study there are codes in Table A.3 that describe the source catalogs in which they were included. Table A.5 is a listing of the source references associated with each case study. Table A.6 documents the magnitude scales used for the earthquakes

tabulated in the catalog (see Section 3.4.1). More detailed descriptions of the various data codes and the data are presented in Appendix A.

3.4 Problems and Details of Catalog Synthesis

3.4.1 Earthquake Magnitude and Distance Measures

Earthquake magnitude measurements can vary by several tenths of a unit, depending on the particular location of the earthquake recording station and the type of recording instrument used. Also, a variety of magnitude scales, including the Richter or local magnitude M_L, body wave magnitude m_b and the surface wave magnitude M_S, have been tabulated in the source liquefaction catalogs. As originally defined, the Richter magnitude M, which seismologists currently refer to as the local magnitude M_L, was measured using the maximum amplitude wave (P, S, or surface wave) recorded by a standard instrument (Wood-Anderson torsion seismograph) which has a specified natural period (0.8 sec), magnification (2800) and damping factor (0.8), corrected to represent a measurement at a standard distance (100 km) from the earthquake epicenter. The local magnitude M_L was intended for use in sourthern California only, being dependent on the attenuation characteristics of that region. In current practice, M_L is evaluated with a variety of seismic recorders in different parts of the world with adjustments for instrument response and regional attentuation characteristics. More details regarding magnitude determination may be found in Bolt (1978) or Aki and Richards (1980).

It has been recognized by seismologists that the local magnitude scale tends to "saturate" for earthquakes with $M_L > 6.5$ or 7.0, i.e. M_L

is actually bounded from above and tends to under-represent the actual size and energy release of large earthquakes. This is due to the combination of the frequency response of seismic instruments and the magnitude-dependent frequency content of the elastic radiation from earthquake sources. Thus in some cases, the magnitude reported for larger earthquakes tend to be the surface wave magnitude M_S, which only "saturates" at a somewhat higher magnitude (M_S > 7.5 or 8.0). Recently, Hanks and Kanamori (1979) have proposed a moment magnitude scale **M**, which does not have problems with saturation. This scale is based on the concept of the "seismic moment", a quantity with units of force times length, which is considered to be more accurate in characterizing earthquake size. However, the measurement of the seismic moment may have as much variation as other size measures. In numerical terms, **M**, M_L and M_S are comparable for magnitudes less than 6.5 to 7.0.

Two additional magnitude scales have been used in the documentation of liquefaction case studies in Japan: the Kawasumi and the Japan Meteorological Agency (JMA) magnitude scales (see Kuribayashi and Tatsuoka, 1975). Estimates of earthquake sizes based on the JMA scale are comparable to those given on the Richter scale, but the Kawasumi magnitude (which is based on an intensity scale) appears to assign slightly larger ratings of earthquake size. The Kawasumi magnitude is usually reported only for historical (pre-instrumental) earthquakes in Japan.

The liquefaction data catalog in Appendix A tabulates the Richter or local magnitude of the earthquake, whenever possible. The rationale is that the local magnitude M_L is believed to have the most relevance to

engineering applications because it is determined based on recorded seismic waves having the frequency range of greatest engineering interest (Kanamori and Jennings, 1978; Idriss, 1978). In many cases the local magnitude M_L has not been measured, or is not available, and in some instances there is even no documentation in the source references regarding which magnitude scale is used. In these latter cases, the only resort is to use the reported magnitude of unknown scale in the compilation. Of the 40 earthquakes reported in Tables A.1 and A.2, 17 are measured on the Richter (M_L) scale, 13 on the JMA scale, 4 on the M_S scale, 3 on the Kawasumi scale, and 3 on an unspecified scale. [See Table A.6 for details.] Due to the various sources of inaccuracy described above, the magnitude reported in the liquefaction catalog should be viewed as only an indicator of earthquake size, and not an exact measurement.

The location of earthquake hypocenters or epicenters and focal depths are determined from arrival times of seismic waves at various recording stations, and are also subject to large variations. For example, in the fairly recent 1979 Montenegro earthquake, the determinations of the epicenter by two different organizations were about 10 km apart (EERI, 1980). Determinations of focal depths and hypocenters tend to have even greater variation.

Another distance measure incorporated in the synthesized catalog is the "distance to energy release" or DER. Where possible DER is defined as the closest distance to the surface fault rupture. In cases where no surface manifestation of the fault rupture is evident, it is defined as the closest distance to the surface projection of the "zone

of energy release", which is sometimes determined from the spatial distribution of earthquake aftershocks. However, if neither of the above measures were available, DER was assumed to be equal to the epicentral distance.

In addition to the inaccuracies of magnitude and distance measures, it should be recognized that earthquakes are often accompanied by foreshocks and aftershocks. In the subsequent inspection by reconnaissance teams, it may not always be easy to discern whether the main shock is solely responsible for the observed liquefaction or other damage. In the synthesized catalog of Appendix A, only the main shock is usually tabulated, but it is conceivable that the build-up of pore water pressures leading to liquefaction could be caused by the cumulative effects of the complex sequence of shock events that accompany an earthquake.

3.4.2 Acceleration and Cyclic Stress Ratio

An important quantity considered in liquefaction analysis is the cyclic stress ratio CSR defined originally by Seed and Idriss (1971) as:

$$CSR = 0.65 \frac{a}{g} \frac{\sigma_v}{\bar{\sigma}_v} r_d \qquad (3.1)$$

where

a = the peak ground surface acceleration

g = the gravitational acceleration (9.81 m/sec^2)

σ_v = the vertical stress at the depth under consideration

$\bar{\sigma}_v$ = the effective vertical stress at the depth under consideration

r_d = a reduction factor that accounts for the soil flexibility as a

fun ion of depth.

In a more recent paper, Seed et al. (1984) have implicitly defined a magnitude normalized CSR as:

$$CSRN = 0.65 \frac{a}{g} \frac{\sigma_v}{\bar{\sigma}_v} \frac{r_d}{r_M} \tag{3.2}$$

where r_M is the magnitude normalization factor. The intent of this normalization is to account for the effects of duration of shaking which is correlated to the earthquake magnitude M. In accordance with Seed et al. (1984), r_M is defined so that CSRN corresponds to CSR for a M = 7.5 earthquake.

The depth reduction factor r_d was derived empirically by Seed and Idriss (1971) based on response analyses of a variety of soil profiles. while the magnitude normalization factor r_M is based largely on laboratory studies and an empirical correlation between earthquake magnitude and duration of shaking (see Seed et al., 1984). Though there is statistical variability associated with r_d and r_M, their variances are negligibly small when compared with the variance of the peak acceleration obtained from an attenuation law or from a nearby measurement. Use of average values of r_d (as a function of depth) and r_M (as a function of magnitude) is recommended in practice, and they are normally given in chart or tabular form (e.g. see Seed and Idriss, 1971 and Seed et al, 1984). For ease of calculation, the following formulas were fitted to the recommended average functions:

$$r_d = \begin{cases} 1.0 - 0.00765\ z & z \leq 9.15\text{m (30 ft)} \\ 1.174 - 0.0276\ z & z > 9.15\text{m (30 ft)} \end{cases} \tag{3.3}$$

where z is the depth in meters;

$$r_M = 0.032\ M^2 - 0.631\ M + 3.934 \qquad (3.4)$$

where M is the earthquake magnitude.

The primary variable affecting the value of CSR or CSRN is the peak ground acceleration 'a', which can be obtained in several ways. In 127 out of 278 catalog entries, the peak acceleration is obtained from measurement at a "nearby" station. The term "nearby" is used loosely here and can mean a strong motion recorder located several kilometers away. In a few cases, a strong motion recorder is actually close enough to be considered "on site", which is also a loosely-defined term indicating that the acceleration measurement is considered to be very accurate in representing the ground motion at the site. However, it is rare to have an accelerometer actually on site. Other methods of estimating acceleration include performing a site response analysis with the input from a ground motion record some distance away, scaled to reflect inferred bedrock motions at the site of interest. In many cases, accelerations are calculated from earthquake attentuation relationships and/or correlations to an intensity damage scale (e.g., Modified Mercalli scale).

In the synthesized catalog presented in Appendix A, the site accelerations tabulated are usually those reported in the source catalogs. However, in many of the historical cases of liquefaction/non-liquefaction from California and Japan where the acceleration was not reported, or where the reported acceleration was suspect, accelerations were estimated from one of two attentuation relationships. For cases in California, the Joyner and Boore (1981) equation was used:

$$\log_{10}(a/g) = -1.02 - 0.249\,\mathbf{M} - \log r - 0.00255\,r \qquad (3.5a)$$

or

$$\frac{a}{g} = \frac{(0.0955)10^{0.249\,\mathbf{M}}}{r \cdot 10^{0.00225r}} \qquad (3.5b)$$

where $\mathbf{M}$ is the moment magnitude, and $r = (d^2 + 7.3^2)^{1/2}$, in which d is defined as the closest distance (in kilometers) to surface projection of the fault rupture. For Japanese earthquakes, the relationship used is that due to Kawashima et al. (1984) for soft alluvium or reclaimed ground:

$$\frac{a}{g} = \frac{0.4109\ 10^{0.262 \cdot M}}{(R_{EP} + 30)^{1.208}} \qquad (3.6)$$

where M is the Japanese Meteorological Association (JMA) magnitude and R_{EP} is the epicentral distance in kilometers. In calculations to estimate accelerations for the synthesized catalog, the tabulated magnitude was used in place of the moment or JMA magnitude.

3.4.3 Correction/Normalization Factors for SPT

There are two corrections or normalizations that need to be made to the N-value obtained directly from the standard penetration test (SPT). The first is to account for the effect of overburden pressure, and the second is to account for the effects of using different sampling equipment and/or practices in performing the SPT. The SPT resistance corrected for overburden is denoted as N_1 and is calculated as:

$$N_1 = C_N \cdot N \tag{3.7}$$

where C_N is the overburden correction/normalization factor. The additional normalization factor to account for sampling equipment and practices is denoted as C_E and the additional correction made is calculated as:

$$(N_1)_{60} = C_E \cdot N_1 = C_E \cdot C_N \cdot N \tag{3.8}$$

The reason for denoting the normalized value as $(N_1)_{60}$ is discussed below.

Several correction factors for overburden have been published in the literature, and a review of these is presented by Liao and Whitman (1986), who propose a simple mathematical form for C_N as:

$$C_N = \sqrt{\frac{1}{\bar{\sigma}_v}} \qquad (\bar{\sigma}_v \text{ in TSF or kg/cm}^2) \tag{3.9}$$

For all practical purposes, this correction factor is equivalent to that proposed by Seed (1979) which is given in chart form. The above equation simply makes the estimation of C_N more convenient and was used in this study.

The values of the correction factor C_E are based on the recommendations of Seed et al. (1984), and are calculated as:

$$C_E = \frac{ER}{60} \cdot C_{JAP} \cdot C_{ROD} \cdot C_{LIN} \tag{3.10}$$

In the above equation, ER is the energy ratio defined as the percent of the theoretical free-fall energy transmitted to the rods from the SPT

hammer. ER has been found to vary depending on the type of hammer used (e.g. "free fall" versus rope and pulley) and on the standards of practice in different parts of the world. The denominator of 60 reflects the recommendation by Seed et al. that SPT data be normalized to an equivalent ER = 60%. C_{JAP} is a correction for the different standards of drilling practice in Japan (i.e. the frequency of hammer drop (blows per minute) and the diameter of the bore hole). C_{ROD} corrects for the effects of short rod lengths when performing the SPT at shallow depths. C_{LIN} is a correction to account for the practice of leaving out the inside liners from the barrel of the SPT sampler, as is done frequently in the U.S. Values of these various correction factors are shown in Table 3.2.

The value of C_E for each case study is tabulated in the liquefaction data catalog in Appendix A. The corrections presented in Table 3.2 are average values, but in specific cases, where the values of ER have actually been measured, these were used instead. In cases where there is uncertainty regarding which type of hammer was used, an average of the possible ER values was employed in calculating C_E, which follows the procedure of Seed et al. (1984).

It should also be noted that there may be disagreement regarding the standard ER value to use in normalization. For example, Kovacs et al. (1984) have proposed that a standard ER = 55% be used instead of ER = 60% as recommended by Seed et al. (1984). For the analyses in this study, the actual standard ER is not considered to be important and would not affect the results (except that the $(N_1)_{60}$ values would all have to be multiplied by a constant to adjust them for a different standard ER).

What is important is that a correction has been attempted so as to make the different types of SPT measurements comparable.

3.4.4 Site Characterization and Data Independence

In the context of liquefaction analysis, site characterization refers to the problem of determining a representative SPT resistance and depth at which liquefaction is likely to occur. In case studies where liquefaction has occurred, it may be possible to identify the depth of liquefaction from comparison of the soil ejected from sand boils with samples obtained at depth in borings. However, such data are not always available, and in the cases of non-liquefaction, an estimation of the critical depth where liquefaction would most likely occur (given a stronger earthquake) can require a considerable degree of judgement.

The related problem of statistical independence of data can be posed as the question: What constitutes a case study data point? Consider the example project site shown in Figure 3.1, which has been reported by Jaime et al. (1981), where liquefaction occurred in a localized zone during the 1979 Guerrero, Mexico Earthquake. In analyzing the situation that occurred, one could treat the boring data in one of three ways:

- Method (1): Consider the borings in the liquefied zone as a single group and the borings in the non-liquefied zone as another group, resulting in a total of two "case studies".

- Method (2): Consider each individual boring as a "case study", resulting for this situation in a total of 34 cases.

- Method (3): Consider each depth in every boring where an SPT

N-value is obtained in sand as a separate "case study", resulting in a total number of data points on the order of hundreds, depending on the frequency of sampling with depth.

Each of the above three methods has been used in compiling previous catalogs. However, even within any given source catalog, no one method has been consistently used, with the result that the catalogs are a hodgepodge in terms of site characterization.

In a situation such as that presented in Fig. 3.1, use of any of the three methods of defining a case study introduces different degrees of correlation along the data. There are two ways to deal with correlation:

(a) Account for correlation in analysis (see Chapter 6 for an example of a procedure to do this), or

(b) Use statistical techniques that assume independence, and eliminate dependent data.

Assuming independence makes analysis easier, and it is the usual assumption in most regression techniques.

Method (1) for defining a case study is preferable if data independence is desired. Since there are two groups of boring locations in Fig. 3.1 that can be distinguished in terms of their different responses to an earthquake, it is reasonable to assume that the soil properties in the two areas are not correlated. However, this approach presents us with the practical problem of site characterization. Certainly we wish to obtain a characteristic N_1 value for each case study, but how? If an average value is desired, over what depths should we obtain an average? Should we characterize the site by the average N_1

value at a given depth from the several borings, or the average of the N_1 values from all depths in each boring? Further complexities can be envisioned as shown in Figure 3.2, where both the depth and the thickness of the liquefiable sand layer is variable across a site. What, in this case, would be a representative depth at which liquefaction would occur, and how would a characteristic N_1 value be obtained?

Method (3), in which every obtained SPT N-value is considered as a "case study", completely eliminates the site characterization problem. In this approach, there is no need to obtain an average or representative N_1-value, since every N_1-value is a data point. In this approach there is an obvious lack of data independence due to the physical proximity of the N-value measurements. In addition, the numerical models of liquefaction point to the view that liquefaction is a propagation phenomenon, in which the excess pore pressures causing liquefaction can initiate within a localized zone, and then propagate by diffusion to other portions of a soil stratum, causing them to liquefy (Seed, Martin and Lysmer, 1976). Considering this view the approach of using each measured N-value as a data point would require a statistical method of analysis that accounts for correlation.

Given the problems of site characterization associated with Method (1) and the lack of statistical data independence inherent in Method (3), the most practical approach is thought to be that of Method (2) where individual borings are used as the natural unit to define a "case study". Using this intermediate approach would introduce a degree of correlation among some data points, but it also allows the development of a simple and consistent method of obtaining a representative N_1-value.

Consider the problem of site characterization exemplified by the hypothetical boring log shown in Figure 3.3. In this profile, there exists a liquefiable layer of soil, but it is not clear whether part or all of the layer would liquefy during a given earthquake. In particular, there seems to be a sub-layer between 5m to 10m depth that is potentially more liquefiable due to its comparatively lower N_1 values.

Simple averages of N_1 over various ranges of depths in the liquefiable soil layer for the hypothetical boring profile are shown in Figure 3.3(b). The range of average N_1 values varies from $(N_1)_{avg} = 6$, if we only use the minimum N_1 value, to $(N_1)_{avg} = 13$, if we average over the entire depth of potentially liquefiable soil. Thus, the representative N_1 value may be highly dependent on one's judgement of the appropriate depths of possible liquefaction occurrence. To remove this element of judgement as a potential source of inconsistencies, the minimum N_1 value is used in this study as the characteristic of a boring profile. The obvious criticism of using this procedure is that the lowest value may actually be a testing error. In such a case, judgement should be used to identify it as such and thus exclude the data point from use in analysis.

It should be noted that the use of the minimum N_1 value is consistent with the concept of liquefaction occurring at a critical depth, as first developed by Seed and Idriss (1971). This is because the critical depth of liquefaction, except in very homogeneous soils, is virtually controlled by the variation of N_1 with depth. Though their methodology was not explicitly stated nor consistently followed, a close examination of the Seed et al. (1975, 1984) catalogs shows that choosing

the minimum N_1 value in each boring or soil profile is in most cases equivalent to their result.

3.4.5 Problems with Yegian and Vitelli Catalog

The Yegian and Vitelli (1981a) catalog of liquefaction cases has made significant use of the approach of obtaining several case study data points from the several N-values obtained in a single boring, i.e., Method (3) described in Section 3.4.4. Their reason for using this approach was based on the concept of evaluating a "point" liquefaction potential, rather than a liquefaction potential representative of the entire soil profile (Yegian, 1984). In some instances, their catalog even goes as far as inferring the depths at which liquefaction did or did not occur in a profile, based on a comparison of paired borings performed before and after an earthquake. If the soil density increased from a looser to a denser state at a given depth (based on N-values), they concluded that this indicated "point" liquefaction at that depth. On the other hand, if the soil became looser or remained at the same density after the earthquake, they designated this as a non-liquefaction data point.

There are several difficulties with the above features of the Yegian and Vitelli (1981a) catalog:

1) If a single observation of a surface manifestation of liquefaction is associated with numerous SPT N-values at various depths, this is not the same as observing liquefaction at several boring locations. The difference between these two types of observations is not recognized in the Yegian and Vitelli catalog.

2) It is dubious whether one can determine at which depths (in a

boring profile) liquefaction did or did not occur, based on changes in SPT resistance alone.

3) The distance between the before/after pair of earthquake borings are frequently not known, and it is not always clear whether they are a proper match.

4) It is likely that in some cases, the rationale for the boring performed before the earthquake was the construction of buildings or other facilities nearby. This brings up the question of how construction activities may have affected the SPT resistances in the post-earthquake boring.

Items (3) and (4) above are also of concern for catalogs other than that of Yegian and Vitelli (1981a). However, it is only in the Yegian and Vitelli catalog that these problems are amplified because of their approach to obtaining data.

3.4.6 Conflicts Between Source Catalogs

In compiling the synthesized catalog, there were several instances where conflicting data were reported for the same case by different source catalogs. Some of the conflicts originated from problems in site characterization, which were discussed earlier in this section. Also, different catalogs may have used the same set of borings, but it was sometimes impossible to identify the correspondence between the case study in one source catalog with that in another. The resolution of such conflicts was accomplished by re-examining the data from the source references to check the judgements and/or calculations of the catalogers. However, in cases where a clear resolution was not possible, the writer deferred judgement to the cataloger who was considered to be more

knowledgeable of the case study, e.g. if the case study was located in Japan, the Tokimatsu and Yoshimi (1983) catalog data took precedence.

3.5 A Profile of the Data

Tables 3.3 and 3.4 and Figures 3.4 through 3.15 are presented as a summary of the liquefaction data catalog. The resulting profile is intended to give a preliminary overview of the data. Important aspects of the data base which affect the results and conclusions derived from them are discussed in subsequent chapters. As shown by the histograms in Figures 3.5 through 3.15, liquefaction occurrences are generally associated with higher measures of earthquake shaking intensity (e.g. CSRN or magnitude) and lower values of soil resistance (e.g. $(N_1)_{60}$).

In anticipation of a significant result of statistical analyses presented in Chapter 4, a set of histograms is shown in Figures 3.15 and 3.17 for a quantity not tabulated in the data catalog, but which can be easily calculated from the data. These are the modified versions of an earthquake load parameter given by Davis and Berrill (1981), which are denoted as Λ_{EP} and Λ_{HY} in this study and defined as:

$$\Lambda_{EP} = \frac{10^{1.5M}}{R_{EP}^{2}\,(\bar{\sigma}_v)^{1.5}} \tag{3.11}$$

and

$$\Lambda_{HY} = \frac{10^{1.5M}}{R_{HY}^{2}\,(\bar{\sigma}_v)^{1.5}} \tag{3.12}$$

where M is the Richter magnitude, $\bar{\sigma}_v$ is the effective vertical stress (in kg/cm^2), R_{EP} is the epicentral distance and R_{HY} is the hypocentral

distance (both in kilometers). The histograms of Fig. 3.16 and 3.17 show normalized version of Λ_{EP} and Λ_{HY} defined as

$$\lambda_{EP} = \Lambda_{EP}/\Lambda_o \tag{3.13}$$

$$\lambda_{HY} = \Lambda_{HY}/\Lambda_o \tag{3.14}$$

where Λ_o is calculated for M = 7.5, R_{EP} or R_{HY} = 100 km, and $\bar{\sigma}_v$ = 1.0 kg/cm^2.

Figures 3.18 and 3.19 show plots of CSRN versus $(N_1)_{60}$ and Λ_{EP} or Λ_{HY} versus $(N_1)_{60}$, also in anticipation of subsequent results. In these figures, the liquefaction data points are plotted as crosses (+) and the non-liquefaction data as open circles (o), which will also be the convention kept for future plots. These plots present a view of the data without any of the lines of equal probability of liquefaction that will be superimposed in subsequent chapters. If the data in Figures 3.18 and 3.19 are compared with similar plots by Seed et al. (1984) or Davis and Berrill (1981), one would note that there is a greater mixing of the liquefaction and non-liquefaction data than there is for these previously published catalogs. It is believed that this greater degree of mixing is the result of a more objective evaluation of boring data than that of Seed et al. (1984) and Davis and Berrill (1981).

3.6 Commentary

This chapter has described the procedures used in the compilation of a catalog of liquefaction occurrences and non-occurrences during earthquakes. A major goal was to try to make the data collection procedures as consistent and objective as possible. The data catalog is

also flexible enough to accommodate future data and possible changes in interpretation of individual case studies.

Considering the various problems encountered in compilation, it is clear that there are potentially several sources of uncertainty in the data. From the determination of earthquake magnitude to the choice of a representative SPT resistance almost all the data include a degree of judgement about the relevant facts of a case study. Virtually none of the data represents a truly objective or "scientific" measurement. However, despite the imperfections, the data represent at this time the best we have in terms of actual empirical links between the various factors that can be measured and observations of field liquefaction (Peck, 1979). Subjecting the data to statistical analyses presented in the following chapters is an attempt to make the best out of an admittedly suboptimal situation.

Table 3.1 - Summary of Liquefaction Source Catalogs

Catalog Code	Source Catalog	No. of Cases	Earthquake Load Parameters	Site Soil Parameters	Comments
1	Whitman (1971)	13	a,CSR,D	z_w, z_L	No indication of soil grain size
2	Seed, Arango and Chan (1975)	38	a,D,M,R,CSR	$M,N_1,D_r,z_w,z_L,\bar{\sigma}_v$	No indication of soil grain size other than broad classifications such as sand, silty sand, etc.
3	Yegian (1976)	58	M,R	$N,N_1,z_w,z_L,\bar{\sigma}_v$	No indication of soil grain size
4	Yegian and Vitelli (1981a)	> 300*	M,R	N,z_n,FC	SPT profiles and grain size curves documented where available. Distance to earthquake source classified as epicentral, hypocentral, or distance to energy release. *Numerous cases based on multiple case studies per boring.
5	Xie (1979)	63	M,R,a,D	z_n,z_L,N	No indication of soil grain size other than broad classifications such as sand, silty sand, etc.

(continued next page)

Table 3.1 - Summary of Liquefaction Source Catalogs
(continued)

Catalog Code	Source Catalog	No. of Cases	Earthquake Load Parameters	Site Soil Parameters	Comments
6	Davis and Berrill (1981)	59	M,R	$\bar{\sigma}_v$,N_1	No indication of soil grain size
7	Tokimato and Yoshimi (1983)	100	a,M,CSR	N,N_1,z_w,z_L,$\bar{\sigma}_v$, FC,CC,GC,D_{50},UC	Liquefaction occurrence classified into 4 states: 1. Extensive liquefaction 2. Moderate liquefaction 3. Marginal site 4. No liquefaction Adjustments made to account for differences in SPT procedures.
8	Seed, Tokimatsu, Harder, and Chung (1984)	125	a,CSR,CSRN,M	N,N_1,$(N_1)_{60}$, D_{50},FC	Adjustments made to account for differences in SPT procedures. Unique feature is $(N_1)_{60}$.

Key:
- a = peak acceleration
- CC = clay content
- CSR = cyclic shear stress ratio
- CSRN = magnitude normalized cyclic stress ratio
- D = earthquake duration
- D_{50} = median grain size
- FC = fines content
- GC = gravel content
- M = earthquake magnitude
- N = SPT resistance, uncorrected
- N_1 = SPT resistance, corrected
- $(N_1)_{60}$ = SPT resistance, corrected and normalized for hammer energy
- R = distance from earthquake source
- UC = uniformity coefficient
- z_L = critical depth of liquefaction
- z_w = depth to water table
- $\bar{\sigma}_v$ = effective vertical stress (at critical depth

Table 3.2

Summary of Factors for Calculating C_E

(After Seed et al., 1984)

Country	Hammer Type	Hammer Release	Estimated Rod Energy ER (%)	Correction Factor for 60% Rod Energy
JAPAN	Donut	Free-Fall	78	78/60 = 1.30
	Donut	Rope & Pulley with special throw release	67	67/60 = 1.12
USA	Safety	Rope & Pulley	60	60/60 = 1.00
	Donut	Rope & Pulley	45	45/60 = 0.75
ARGENTINA	Donut	Rope & Pulley	45	45/60 = 0.75
CHINA	Donut	Free-Fall	60	60/60 = 1.00
	Donut	Rope & Pulley	50	50/60 = 0.83

Additional Factors:

C_{JAP} = 0.9 for Japanese data, $(N_1)_{60} < 20$.

C_{ROD} = 0.75 for depths of SPT less than 3.0 m (10 ft).

C_{LIN} = 1.1 to 1.2 if inside diameter of SPT is not constant, i.e no liners should be used for loose sands and 1.2 should be for dense sands.

For conditions other than specified above, use C_{JAP}, C_{ROD}, or C_{LIN} equal to 1.0.

Table 3.3

Profile and Inventory of the Data Catalog

ITEM/CATEGORIES	No. of Cases LIQ.	NO LIQ.	TOTAL
Site Performance			
Liquefaction	-	-	114
No liquefaction	-	-	164
Geographic Location			
Japan	51	69	120
California	28	72	100
China	15	25	20
Other	20	18	38
Epicentral Distance			
Based on Instrumental Epicenter	70.	.94	164
Based on Location of Maximum Intensity	44.	.70	114
Distance to Energy Release			
Assumed same as Epicentral Distance	64.	106	170
Estimated Independently of Epicentral Distance	50.	58	108
Acceleration Estimate			
Strong Motion Recorder "On" Site	5	19	24
Strong Motion Recorder "Nearby"	.49	78	127
Site Response Analysis Performed	5	11	16
Based on Attenuation Relationships in Section	.15	46	61
Based on Other Attenuation/Intensity Data	.41	17	58.
N-Value Measurement			
SPT - Rope and Pulley	32	43	75
SPT - Free Fall Hammer	38	41	79
SPT - Method Unknown	38	69	107
Estimated From Static Cone	1	0	1
Based on Other than SPT or Static Cone	15	11	16

Table 3.3 (Continued)

ITEM/CATEGORIES	No. of Cases LIQ.	NO LIQ.	TOTAL
Method of Obtaining Representative N-Value			
Critical or Minimum N_1-Value	55	96	151
Layer Average (One Boring)	4	2	6
Site Average (Many Borings)	19	26	45
Method Unknow or Arbitrary	36	40	76
Level Ground			
Slopes, Dams, Dikes, Embankments	7	3	10
Level Ground (Includes Cases with Building Foundations Nearby)	107	161	268
Ground Water Conditions			
Artesian Conditions Noted	2	0	2
Artesian Conditions Not Noted	112	64	276
Time Boring Performed			
Before Eaerthquake	12	27	39
After EArthquake	83	94	177
Unknown	19	43	62
Soil Gradation			
Very Fine Sands ($D_{50} \leq 0.25$)	45	49	94
Not Very Fine Sands ($D_{50} > 0.25$)	69	115	184
Clean Sands[1] (< 12% Fines)	68	114	182
Silty Sands[1] (> 12% Fines)	46	50	96
Non-Gravely Sands (< 10% Gravel)	106	160	266
Gravely Sands (> 10% Gravel)	8	4	12
Uniform Sands (UC < 6)	95	147	242
Well-Graded Sands (UC > 6)	19	17	36

Note:

1. Where fines content data are not available, it is assumed that the sand can be classified as "clean", unless otherwise indicated by verbal descriptions or classifications.

Table 3.4

Summary of Missing Data Items
in the Liquefaction Catalog

Data Item	No. of Cases where Item is Available	Missing
Earthquake Focal Depth	228	50
Duration of Earthquake Shaking	101	177
Fines Content	159	119
Clay Content[1]	137	141
Gravel Content[2]	164	114
D_{50} Size	165	113
Uniformity Coefficient	124	154
All Gradation Attributes	113	165

Notes:

1. Of the 137 cases where clay content data are available, 97 cases have clay content = 0. Only 40 cases have non-zero clay content.

2. Of the 164 cases where gravel content data are available, 129 cases have gravel content = 0. Only 35 cases have non-zero gravel content.

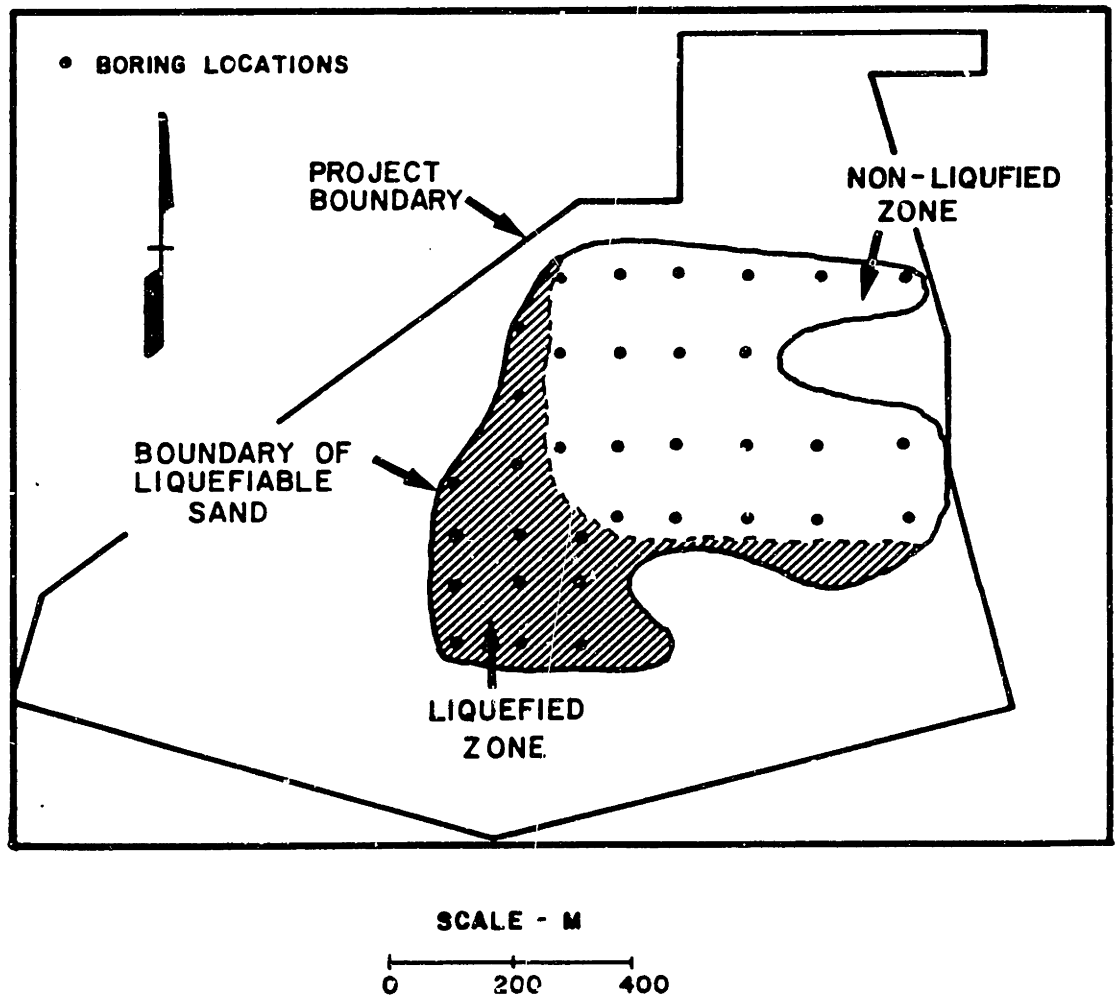


Fig. 3.1 Example project site for determination of case study data. After Jaime et al. (1981).

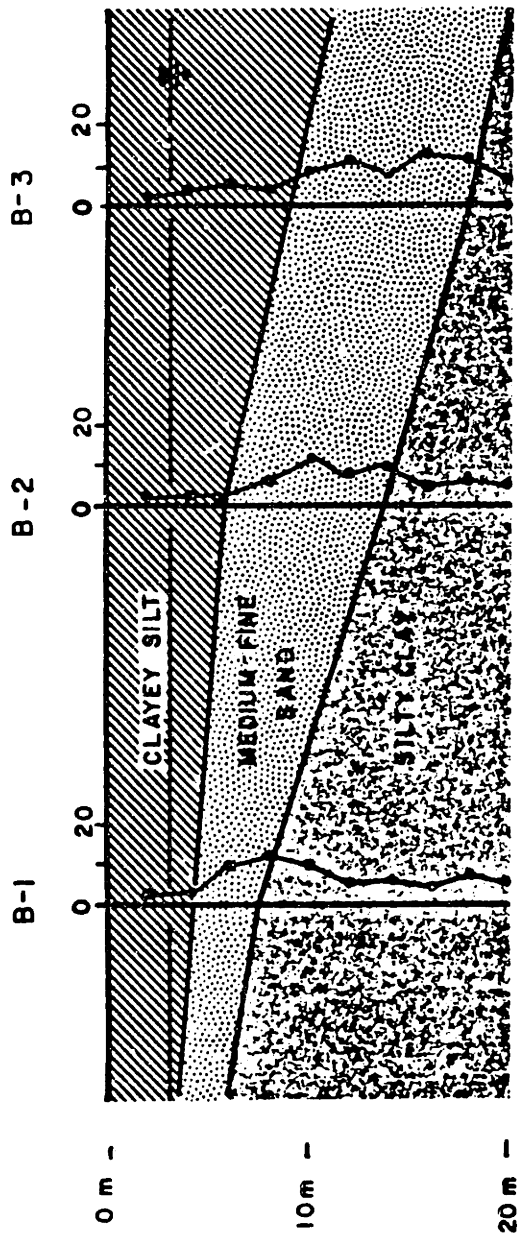


Fig. 3.2 Hypotheitcal soil profile illustrating difficulties of site characterization using several borings.

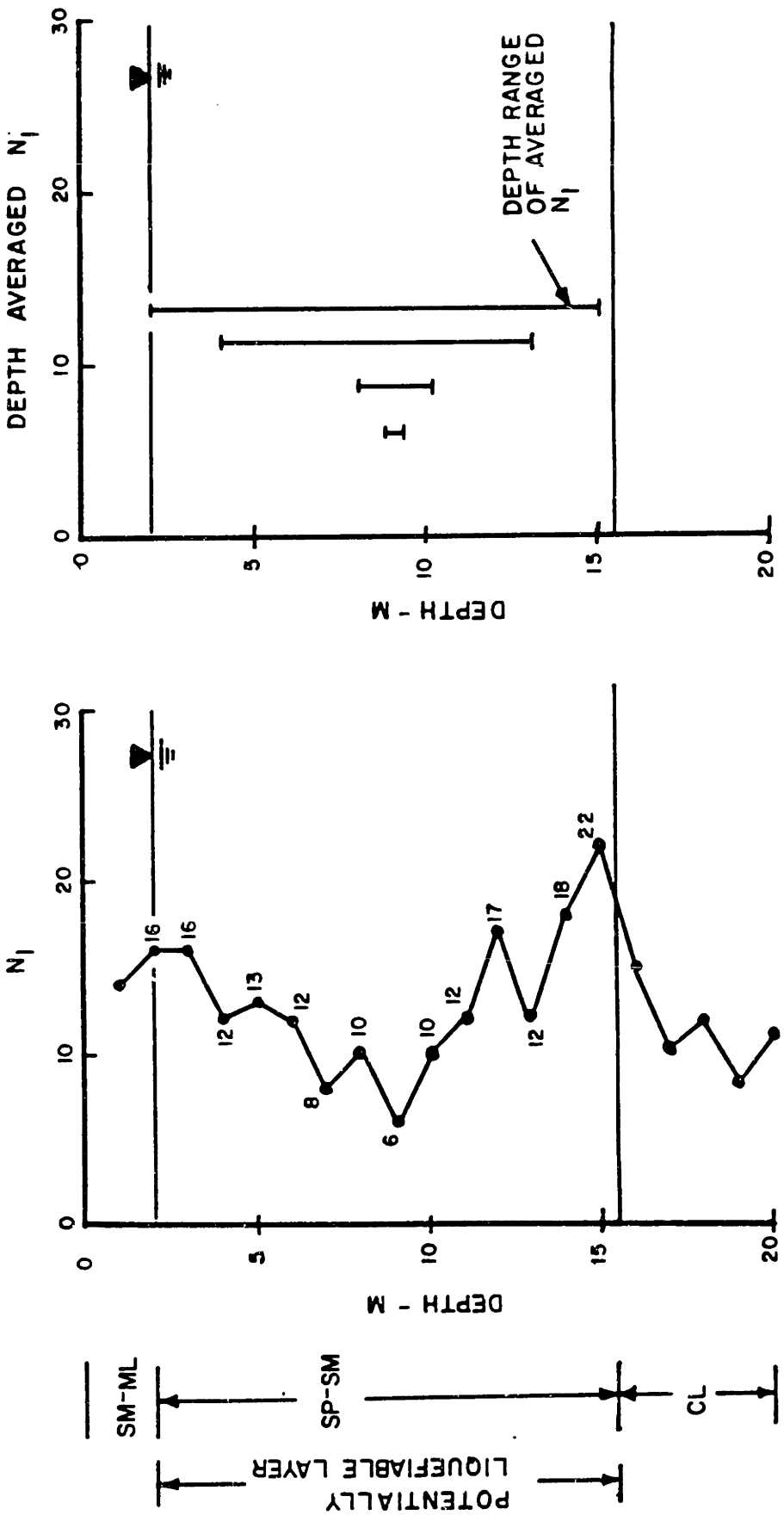
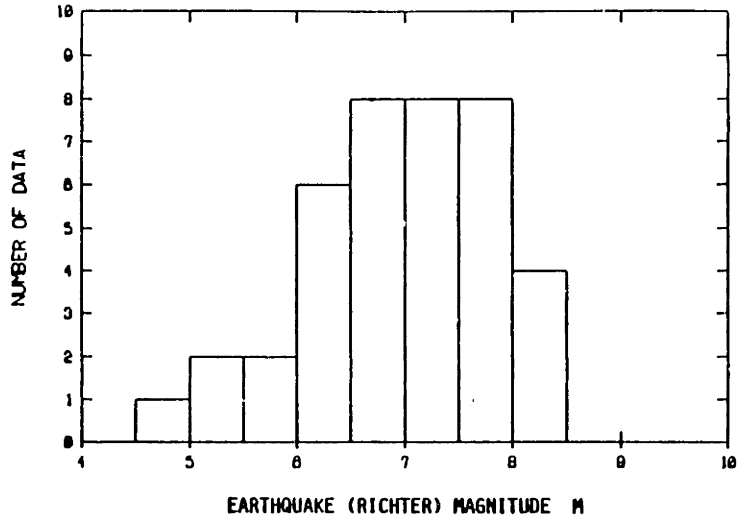
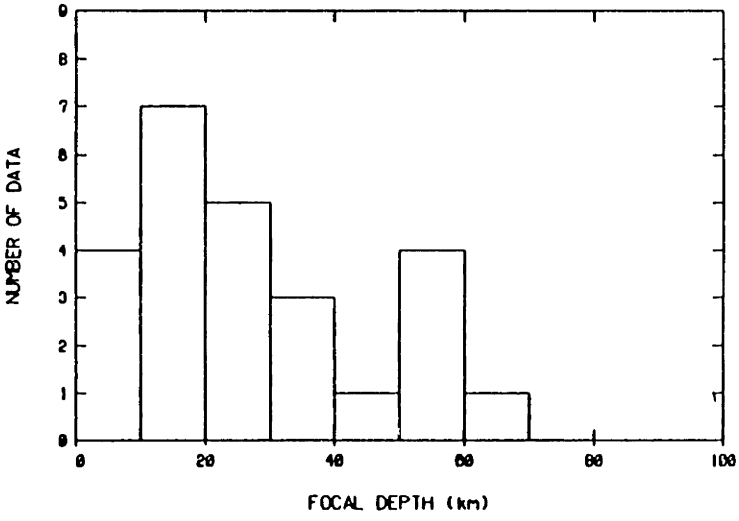


Fig. 3.3 Hypothetical boring profile illustrating potential site characterization problems for a single boring.

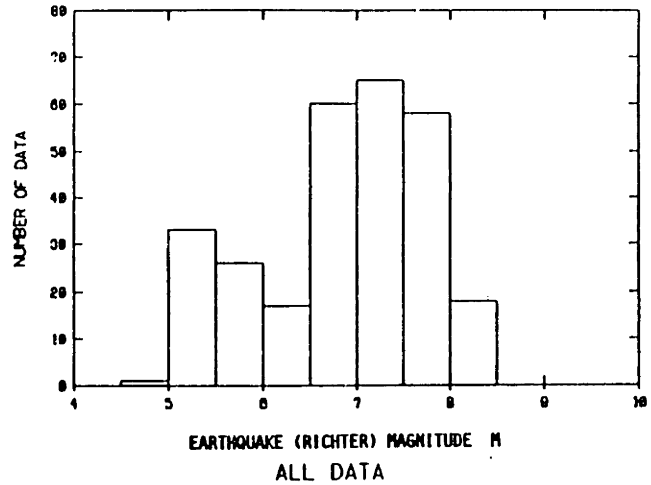


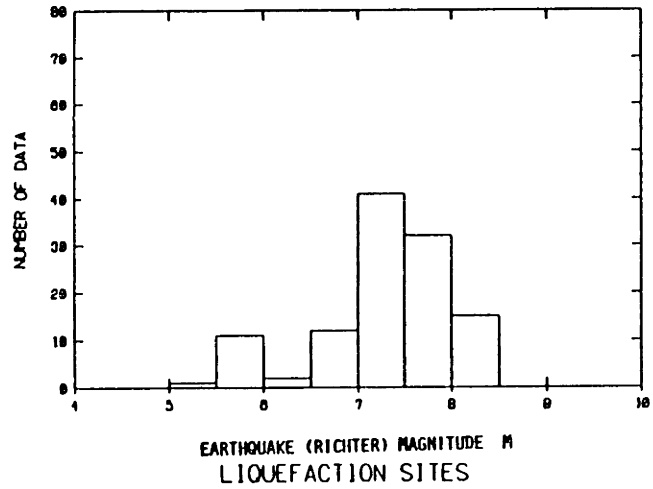
(a) Histogram for the 40 earthquakes in the edited/enhanced catalog.



(b) Histogram for the 25 earthquakes with focal depth data in the edited/enhanced catalog.

Fig. 3.4 Histograms of earthquake magnitudes and focal depths





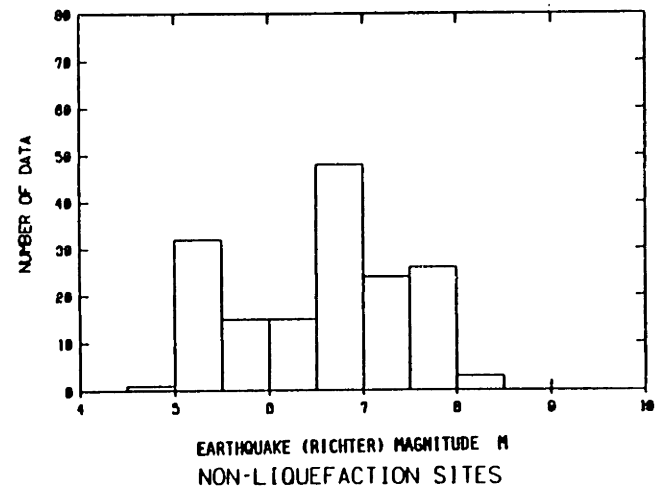
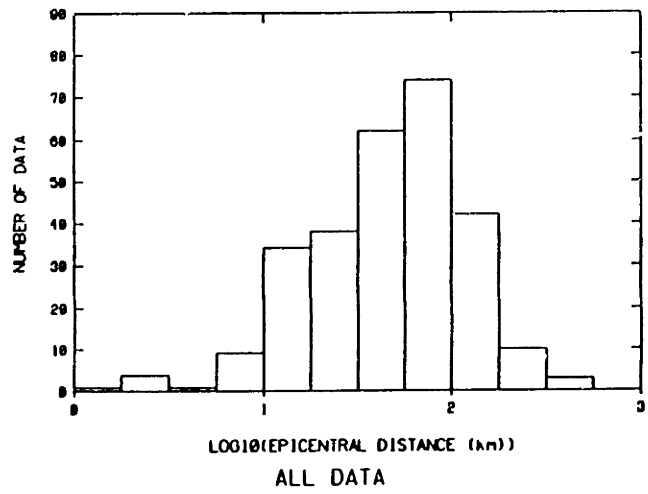
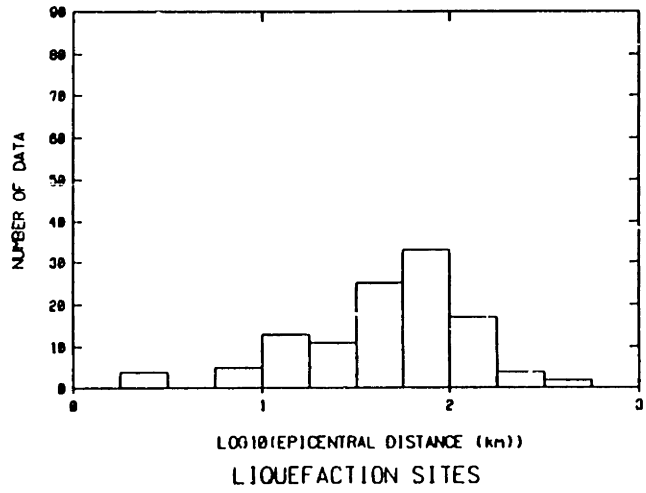


Fig. 3.5 Histograms of earthquake magnitudes associated with each of 278 cases (114 liquefaction, 164 non-liquefaction). Note: several cases can arise from one earthquake.





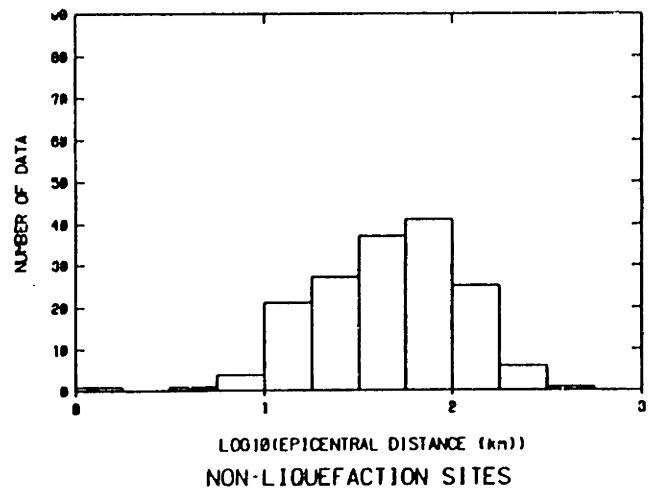
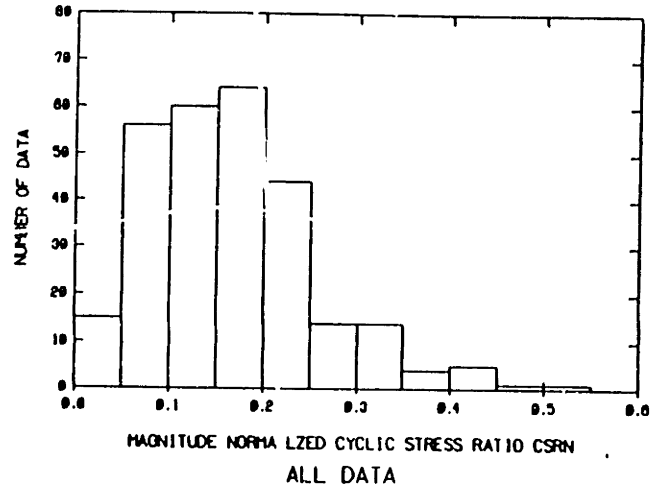
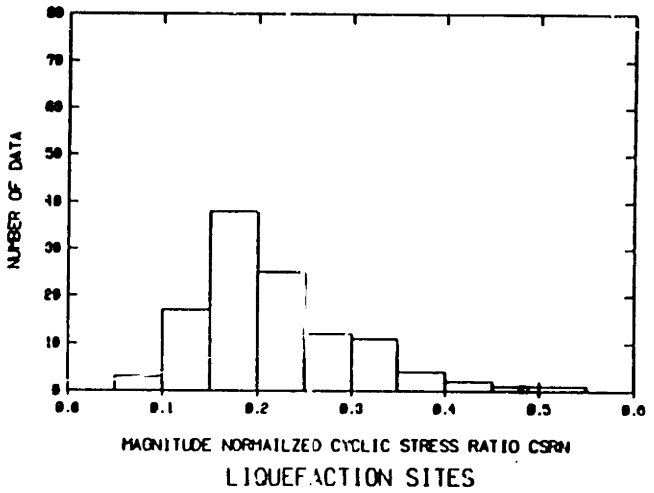


Fig. 3.6 Histograms of $\log_{10}$ of the epicentral distance in km for 278 cases (114 liquefaction, 164 non-liquefaction).





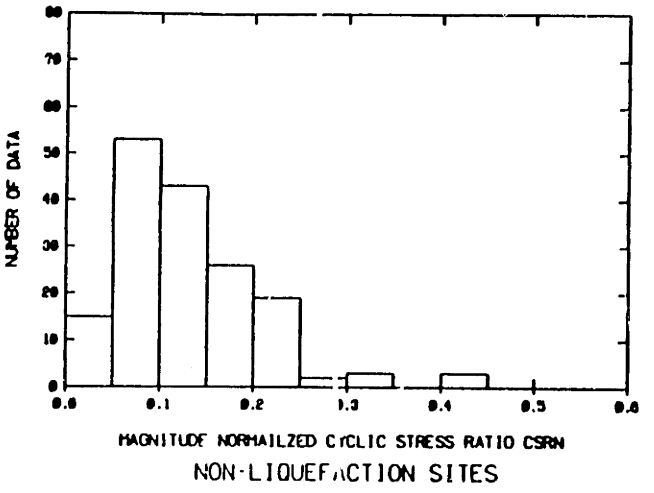
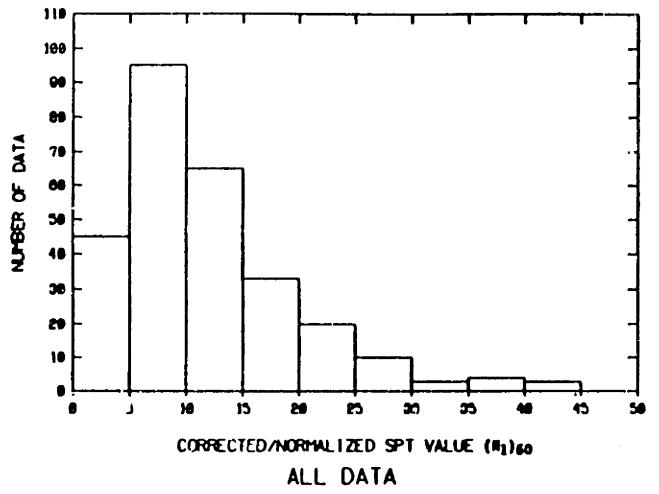
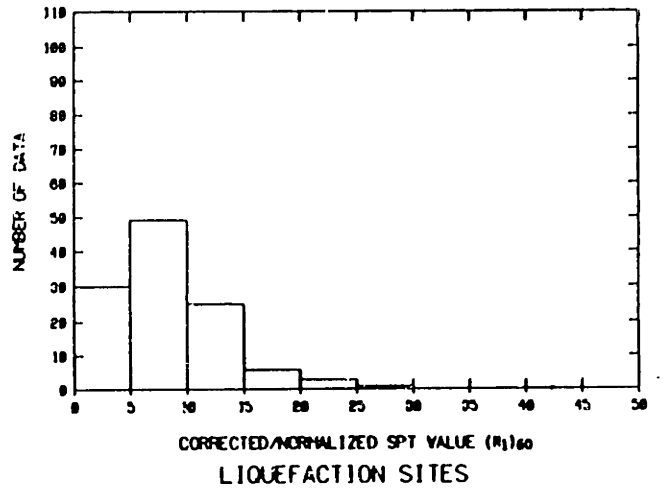


Fig. 3.7 Histograms of the magnitude normalized cyclic stress ratio CSRN for 278 cases (114 liquefaction, 164 non-liquefaction).





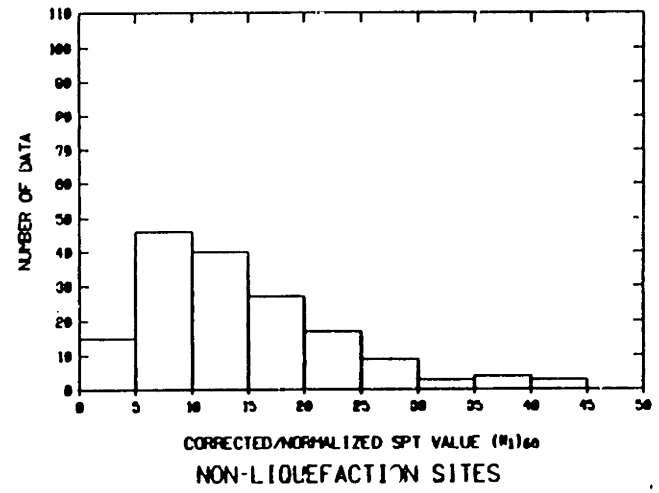
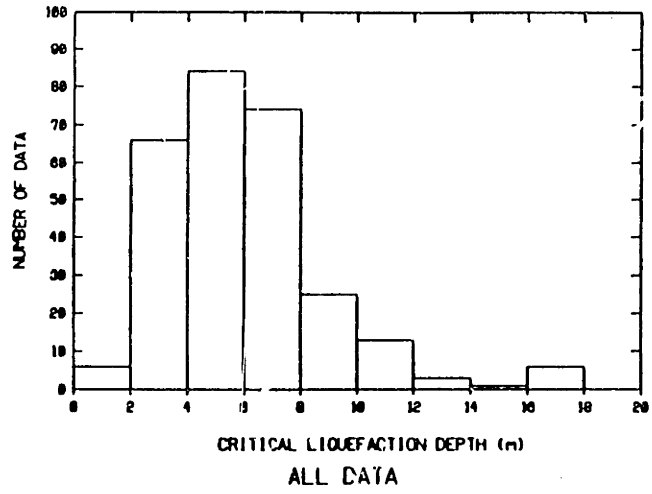
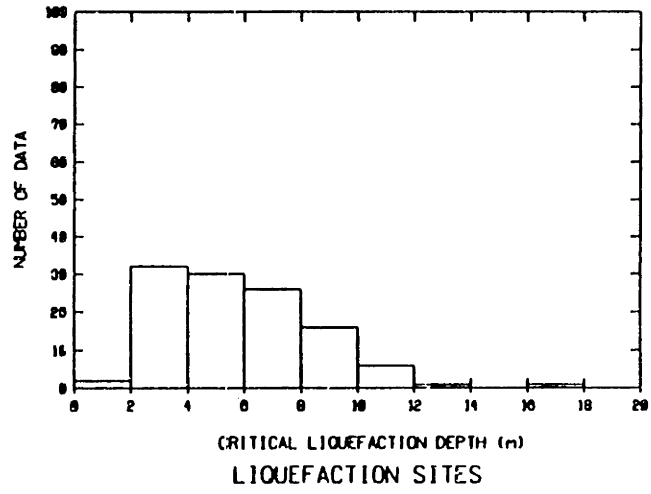


Fig. 3.8 Histograms of the corrected/normalities SPT value $(N_1)_{60}$ for 278 cases (114 liquefaction, 164 non-liquefaction).





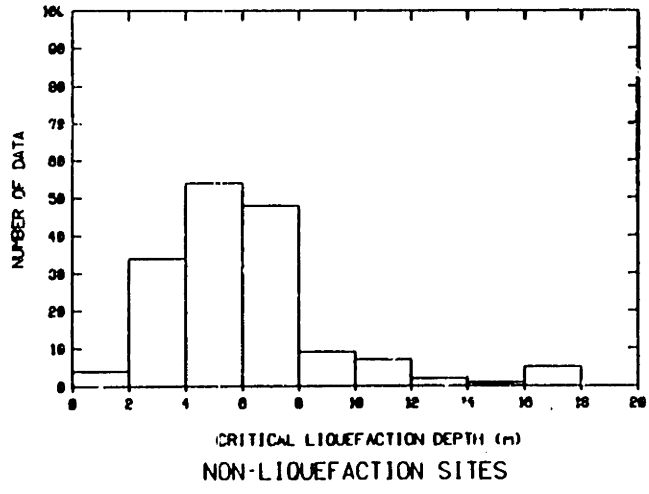


Fig. 3.9 Histograms of the critical liquefaction depth in meters for 278 cases (114 liquefaction, 164 non-liquefaction).

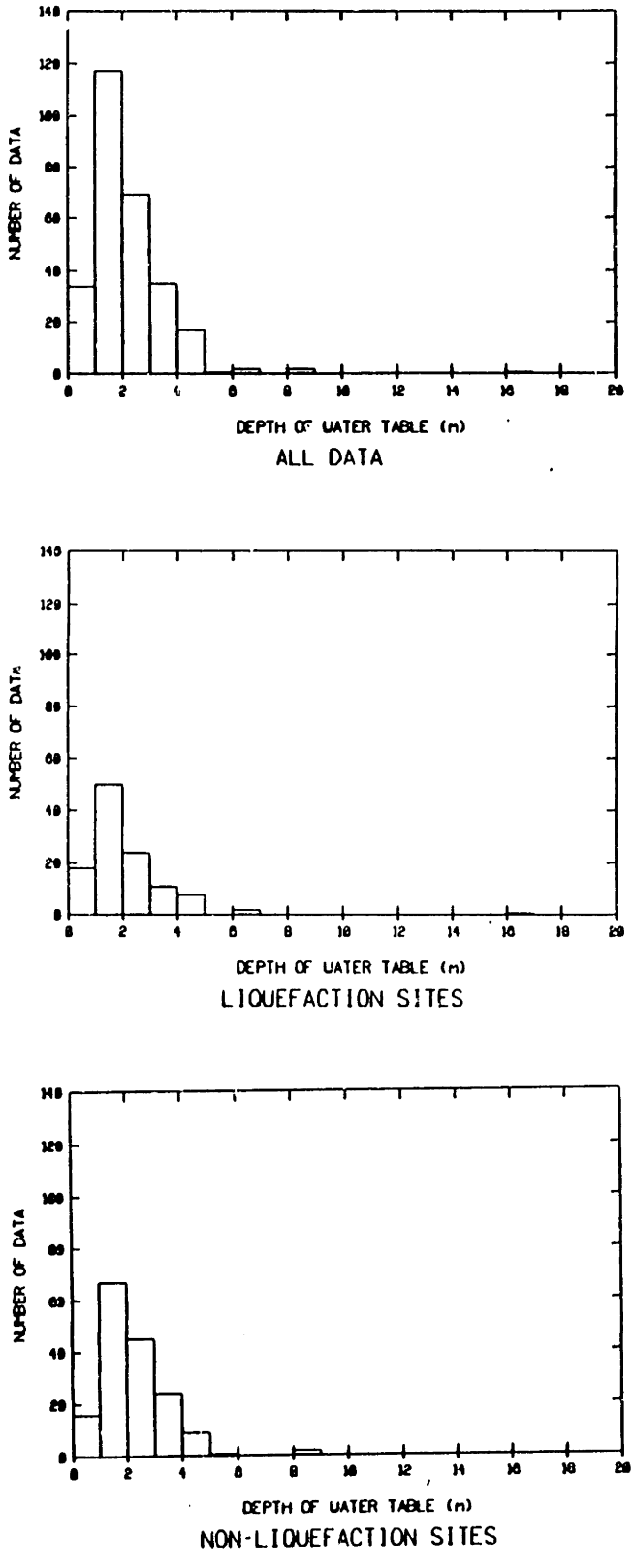
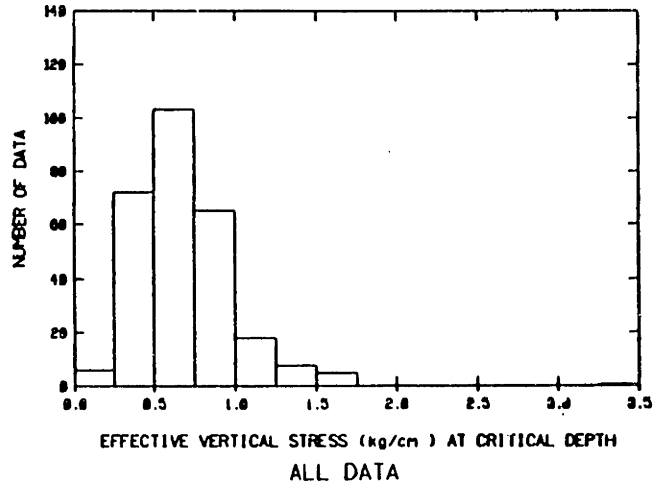
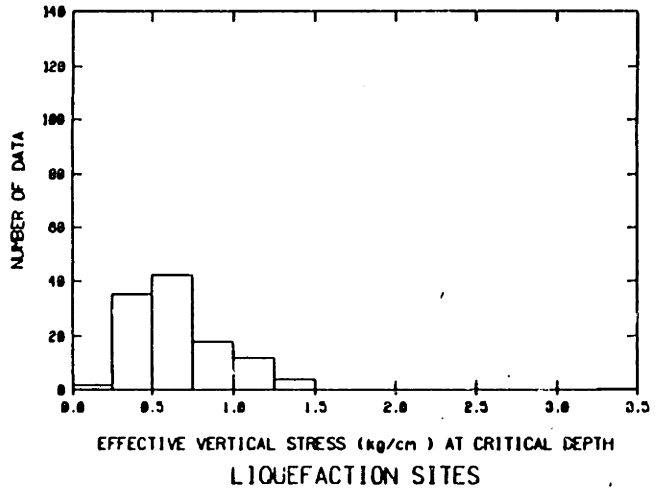


Fig. 3.10 Histograms of depth of water table in meters for 278 cases (114 liquefaction, 164 non-liquefaction).





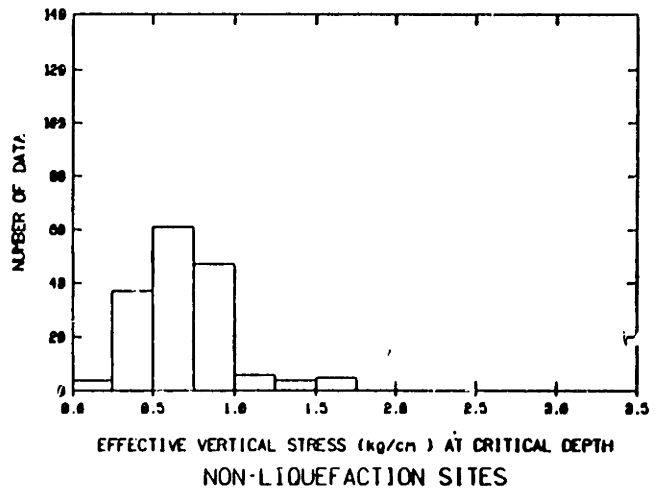
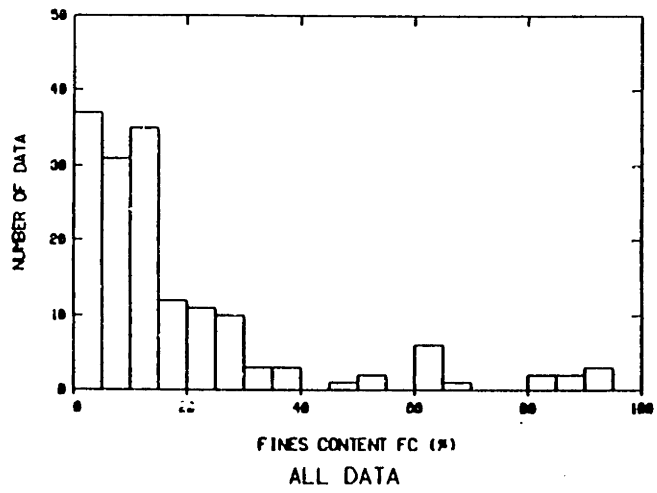
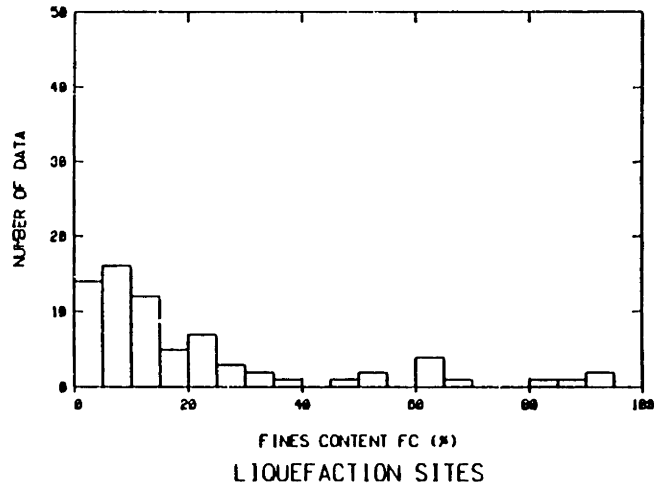


Fig. 3.11 Histograms of effective vertical stress in kg/cm^2 at the critical liquefaction depth for 278 cases (114 liquefaction, 164 non-liquefaction).





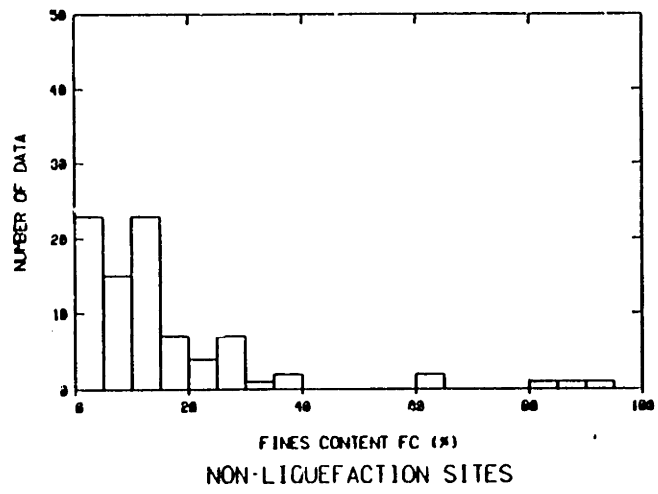
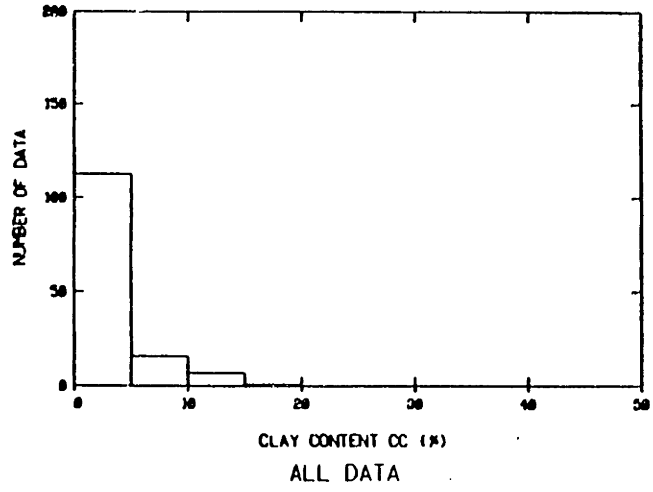
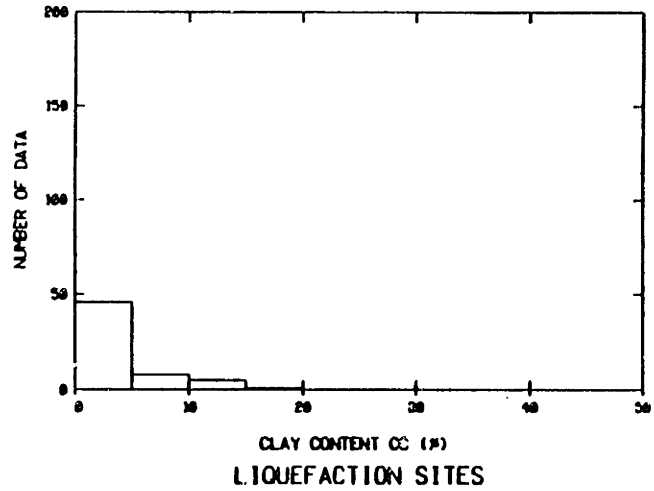


Fig. 3.12 Histograms of fines content FC in percent for 159 cases (72 liquefaction, 87 non-liquefaction).





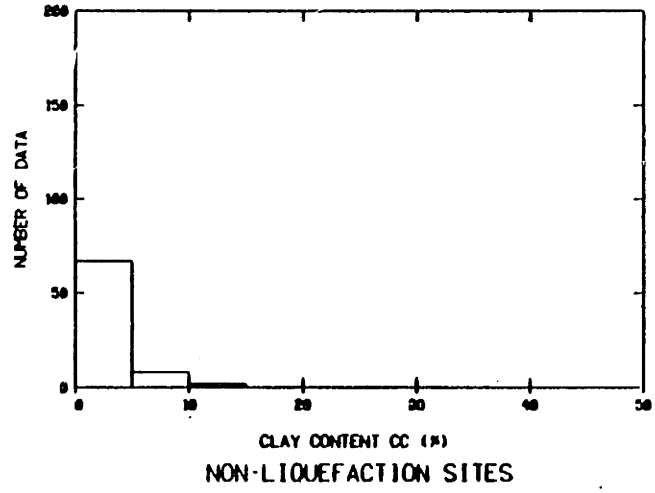


Fig. 3.13 Histograms of clay content CC in percent for 137 cases (60 liquefaction, 77 non-liquefaction).

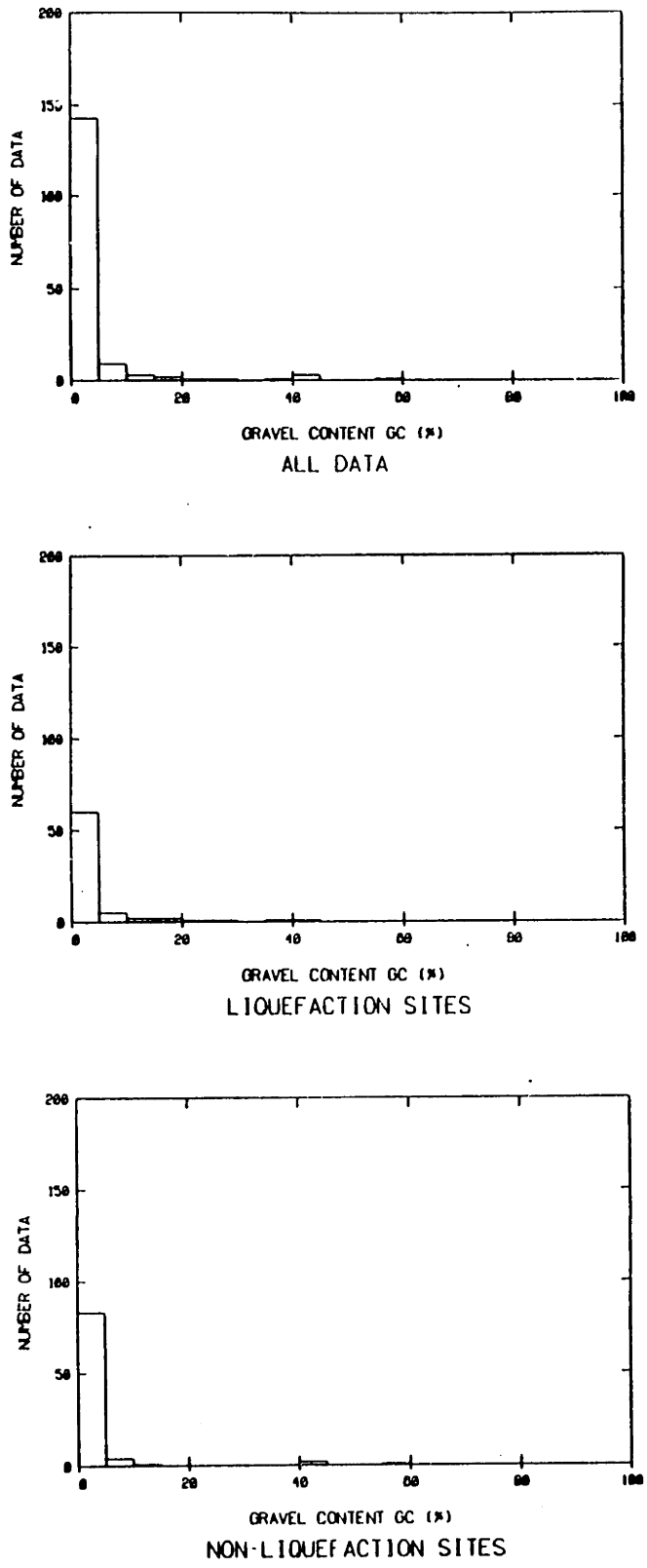


Fig. 3.14 Histograms of gravel content GC in percent for 164 cases (73 liquefaction, 91 non-liquefaction).

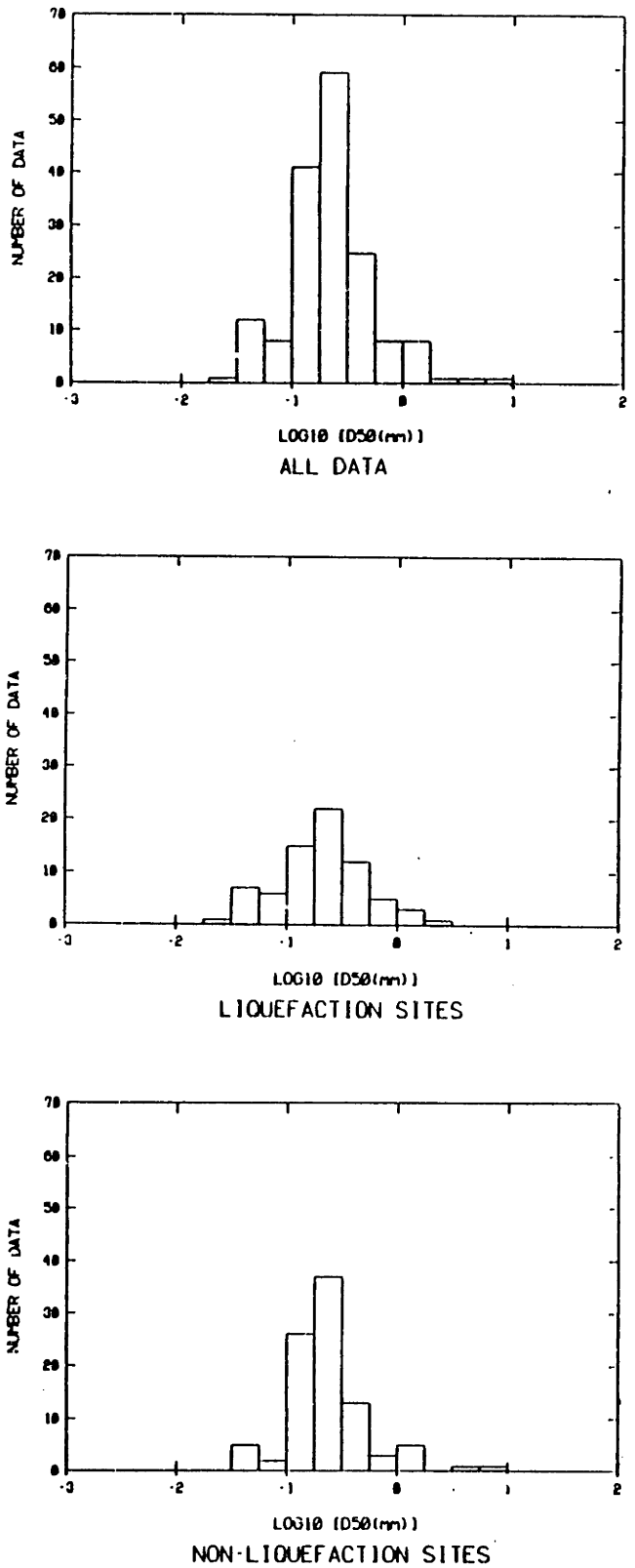
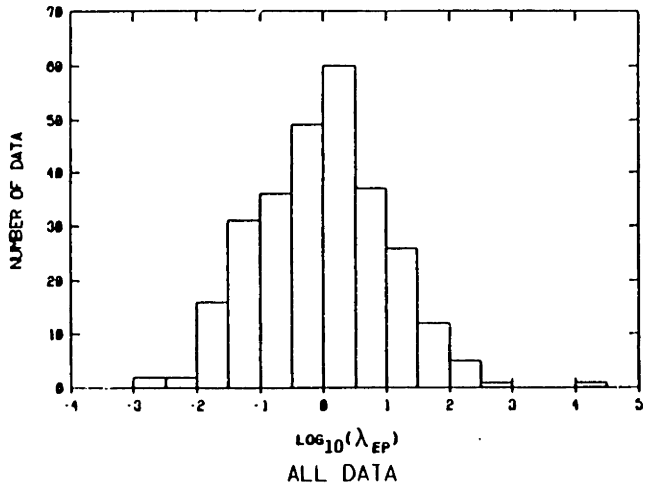
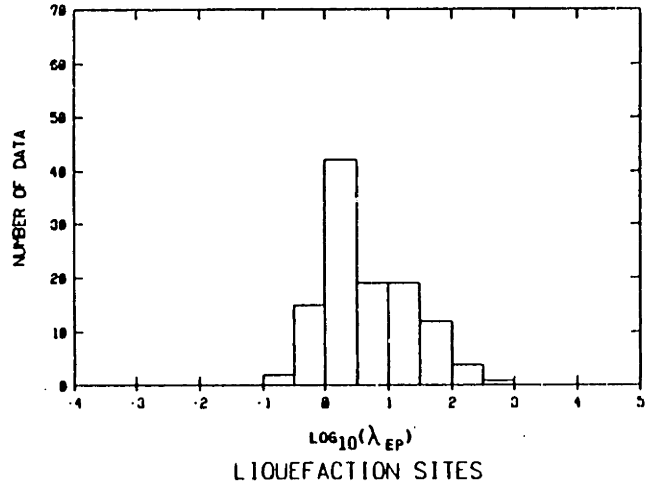


Fig. 3.15 Histograms of $\log_{10}$ of the median grain size D_{50} in mm for 165 cases (72 liquefaction, 93 non-liquefaction)





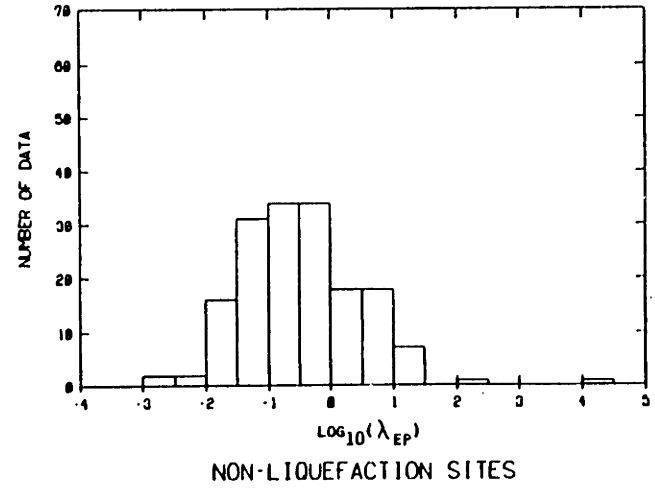


Fig. 3.16 Histograms of Davis and Berrill (1981) normalized load parameter λ_{EP} for 278 cases (114 liquefaction, 164 non-liquefaction)

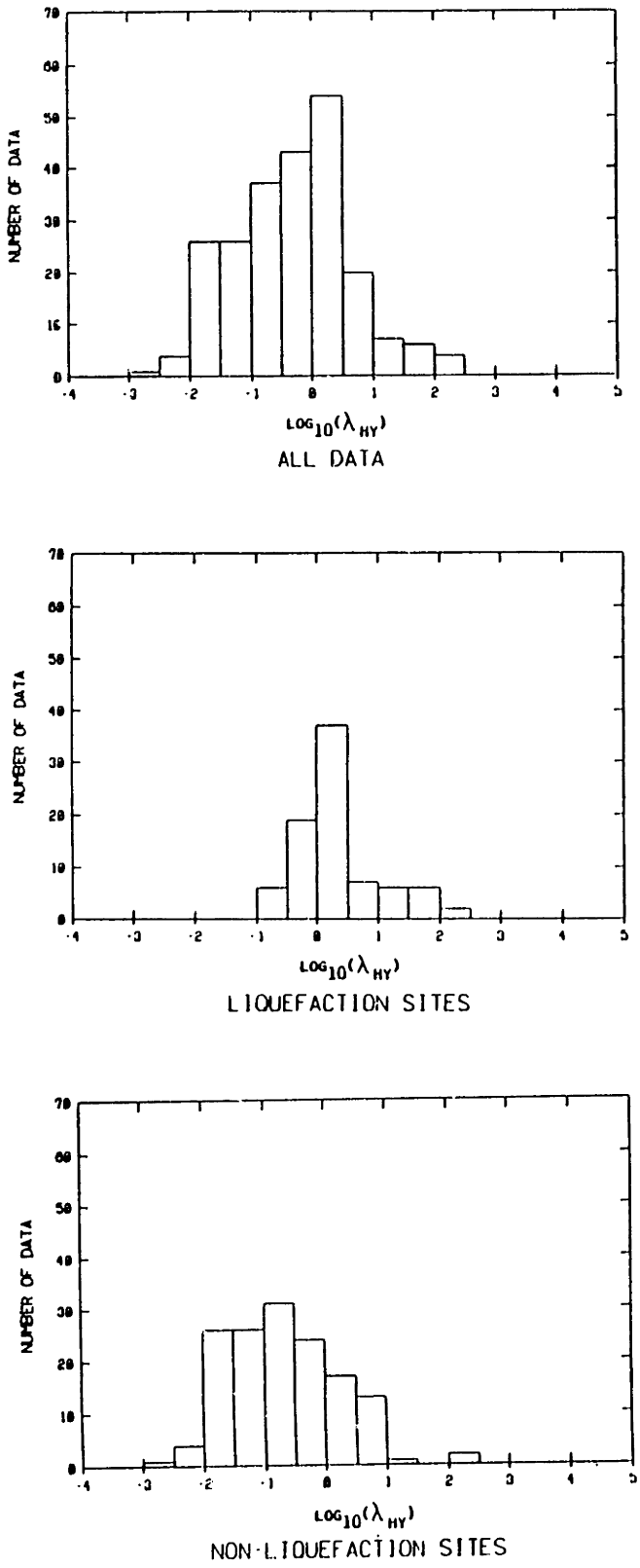
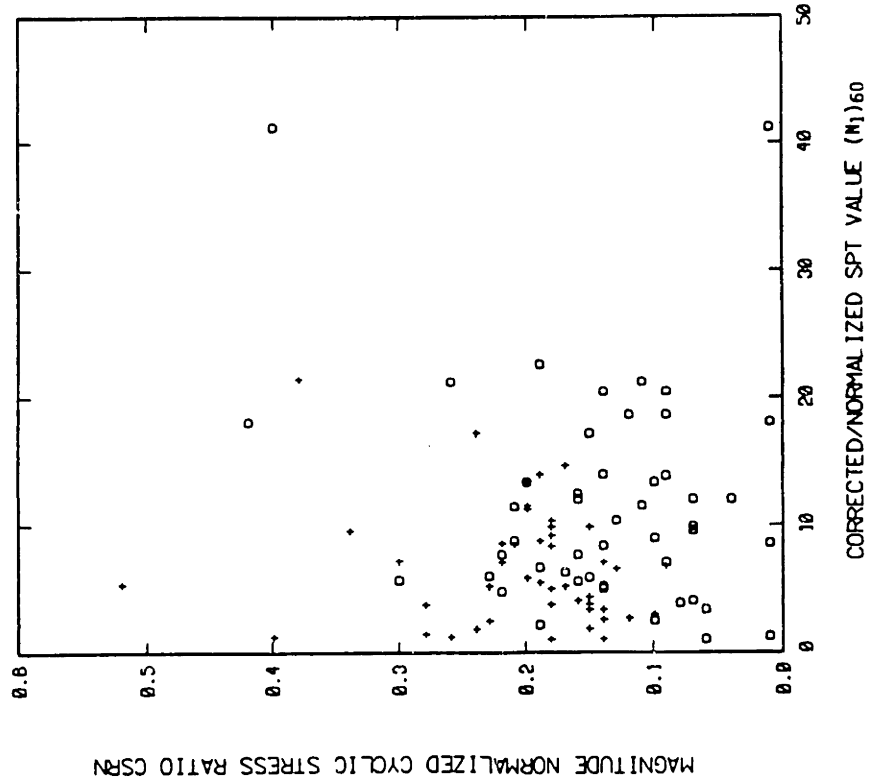
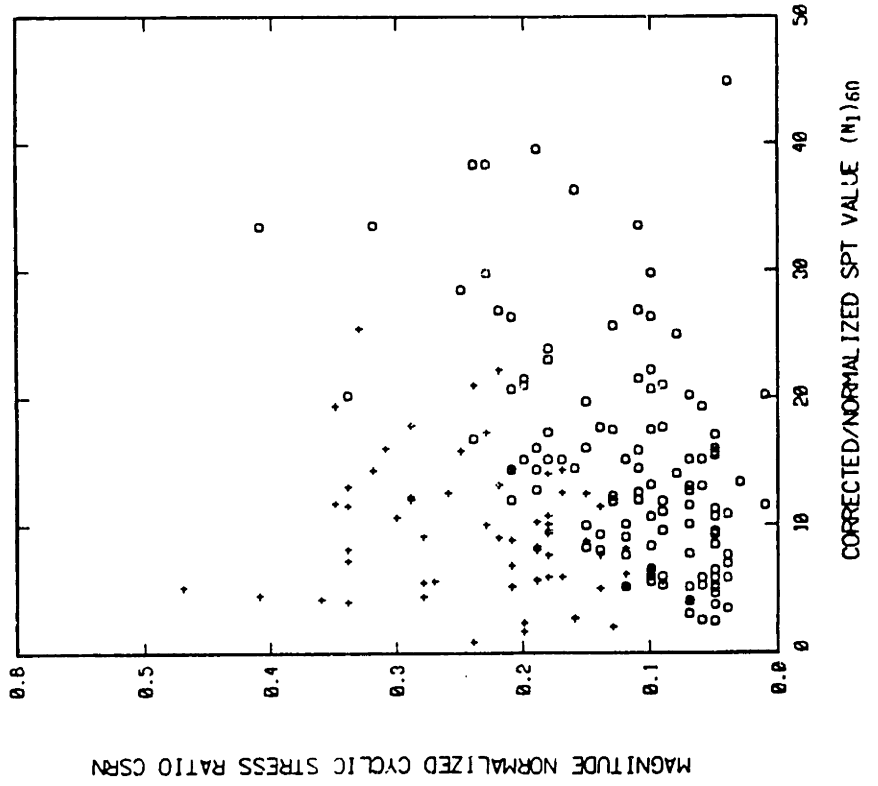


Fig. 3.17 Histograms of Davis and Berrill (1981) normalized load parameter λ_{HY} for 228 cases with focal depth data (83 liquefaction, 145 non-liquefaction)

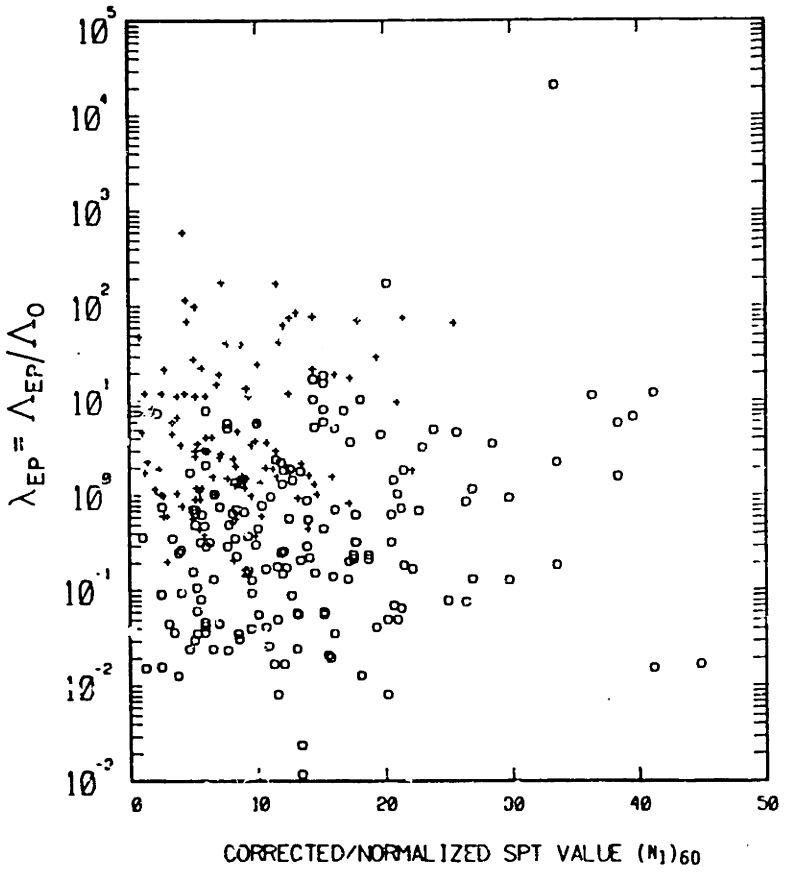


(b) Silty sand data

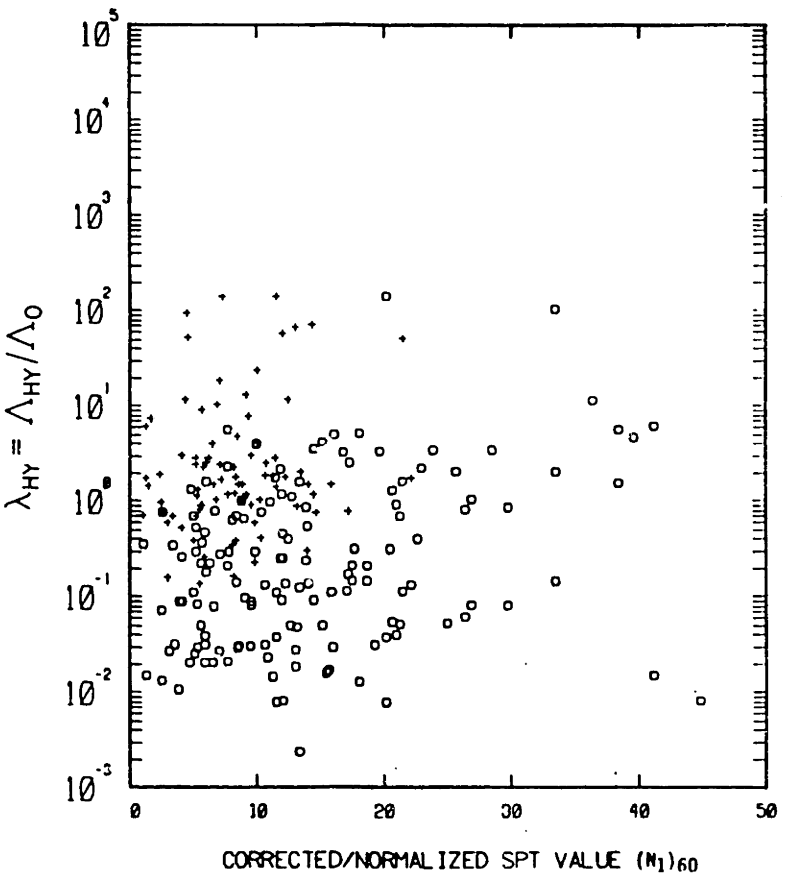


(a) Clean sand data

Fig. 3.18 Data plots of magnitude normalized cyclic stress ratio versus $(N_1)60$. Note: (+) indicates liquefaction, (o) indicates no liquefaction.



(a) Epicentral model



(b) Hypocentral model

Fig. 3.19 Data plots of modified versions of Davis and Berrill (1981) normalized load parameters versus $(N_1)_{60}$. Note: (+) indicates liquefaction, (o) indicates non-liquefaction.

CHAPTER 4

BINARY LOGIT ANALYSIS - METHOD AND RESULTS

4.1 Introduction

Logit analysis or logistic regression is a very powerful and flexible tool for analyzing liquefaction data. Essentially, the objective of logit analysis is to obtain an expression for the conditional probability of liquefaction P as:

$$P = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 \ldots + \beta_m x_m)} \qquad (4.1a)$$

or equivalently:

$$P = 1/[1 + \exp\{-(\beta_0 + \beta_1 x_2 + \beta_2 x_2 + \ldots + \beta_m x_m)\}] \qquad (4.1b)$$

In the above equation, the x_k's (k=1,2,...m) are various "explanatory variables" such as cyclic stress ratio and corrected SPT resistance. Each combination of various explanatory variables is referred to here as a "model". The β_k's (k=0,1,...m) are regression coefficients to be obtained by fitting Equation 4.1 to data.

It is emphasized that logit analysis is a method of regression, but in contrast to standard regression analyses, the response variable Y is "dichotomous" or "binary", e.g. Y=1 if liquefaction occurs and Y=0 if it does not. In formalized notation, the technique works by obtaining the expectation that Y=1, given x_1, x_2 ... x_m, i.e.:

$$P \equiv P(Y=1 \mid x_1, x_2, \ldots x_m]) = E[Y=1 \mid x_1, x_2, \ldots x_m] \qquad (4.2)$$

where E denotes expectation

In Section 4.2, the basis of logit analysis is presented showing the reason for its appropriateness in the analysis of liquefaction data. The maximum likelihood method for estimating the β coefficients is briefly outlined in Section 4.3 along with other aspects of implementation of logit regression. Subsequent sections of this chapter discuss the analyses of the data, focusing primarily on the question of what is the "best" set of explanatory variables x_k to be included in Equation 4.1.

4.2 Logit Analysis: Methodology

4.2.1 Basic Concept

Logit analysis or logistic regression has its origins in biological assay (Berkson, 1944). It was preceded in its development by another procedure called probit analysis, a competing method which is also frequently used. "Probit" was a term coined by Bliss (1934a, 1934b) as a contraction for "probability unit" and "logit" was similarly baptized by Berkson (1944) as a contraction for "logistic unit". "Logistic" refers to the probability function employed in the analysis, and which was originally used as a mathematical model in studies of population growth (Verhulst, 1845; Pearl and Reed, 1920). More contemporary references on logistic regression include those by Cox (1970), McFadden (1974), Dobson (1983) and Ben-Akiva and Lerman (1985). The term "logit analysis" is commonly used by workers in

transportation and consumer demand modelling, whereas "logistic regression" is preferred by statisticians. Both terms are used interchangeably in this thesis.

The original application of probit and logit analysis in biological assay is that of estimating dose-response curves, and can be best explained by an example. Suppose we are interested in evaluating the effectiveness of a chemical to be used as an insecticide. To do this, a large number of insects are taken and divided randomly into several batches with approximately the same number of insects in each batch, e.g. 1000 insects divided into 5 groups of 200. Then each of the batches is administered a different dosage of the insecticide, and the responses of each insect -- death or survival -- is noted. In each batch, a certain percentage of the insects are killed, and this percentage increases with the log of the dosage, e.g. as shown in Fig. 4.1(a).

At any given dosage, all the insects do not have the same response because of natural variations in their individual immunities. This unmeasured, or perhaps, unmeasurable immunity gives rise to the uncertainty about the response of any individual insect. Thus there are two components to any individual response: that which is "explained" by the dosage administered, and that which is unexplained and considered random.

The fraction killed, p, can be transformed to logistic units or logits using the following formula:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) \tag{4.3}$$

Another common name for the above transform is the log-odds ratio, i.e.

the logarithm of the chance or odds of death to survivial. Notice that as p ranges from 0 to 1, the logits of p ranges from $-\infty$ to $+\infty$. If the logit values of p are plotted against the log dosage, the data points tend to fall on a relatively straight line, as shown in Fig. 4.1(b). A similar plot can be obtained using probits of p defined as:

$$\text{probit}(p) = \Phi^{-1}(p) \tag{4.4}$$

where Φ^{-1} denotes the inverse standard normal cumulative distribution function. [In some applications, the probit transform is defined with an additive constant of 5 units, i.e. $\text{probit}(p) = \Phi^{-1}(p) + 5$, e.g. see Finney (1971).]

In either case, whether one uses probits or logits, a straight line can then be fitted to the data. For logit analysis, the equation is of the form:

$$\text{logit}(P) = \beta_0 + \beta_1 x \tag{4.5a}$$

or

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 x \tag{4.5b}$$

where x, in this example, is the log of the insecticide dosage and β_0 and β_1 are the intercept and slope coefficients to be determined from the data. A change in notation from lower case 'p' to capital 'P' is effected to distinguish between the measured fractions p from P fitted to the data. The line represented by Eqn. 4.5 can be roughly fitted by eye, as was originally done by Bliss (1934a, 1934b) or by methods of least squares regression or maximum likelihood. The quantity P can be

interpreted as an estimate of the conditional probability of death for any individual insect, given the log dosage x of the pesticide. If Equation 4.5 is solved for P, one obtains:

$$P = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)} = \frac{1}{1 + \exp\{-(\beta_0 + \beta_1 x)\}} \tag{4.6}$$

The above equation is of the form of a logistic cumulative distribution function.

4.2.2 Generalizations of the Basic Concept

There are three generalizations of logit regression as just described that greatly enhance its usefulness and flexibility. The first is the consideration of more than one explanatory variable x that influences the outcome or the response. For example, in addition to the dosage of insecticide administered, the size of the insects might also be a determinant of the outcome, and we could divide the insects into various sizes groups within each dosage batch. In terms of the mathematics involved, the equation to be fitted would now become:

$$\text{logit}(P) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \tag{4.7}$$

where x_1 is the log dosage and x_2 is the insect size. Further generalization to m variables is straightforward with equations of the type:

$$\text{logit}(P) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m \tag{4.8}$$

or, as often written in vector notation:

$$\text{logit}(P) = \underline{x}^T \underline{\beta} \tag{4.9}$$

where

$$\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_m \end{bmatrix} \qquad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}$$

The second generalization involves the nature of the explanatory variables x_k. In logit analysis, there is no reason why one should restrict the x_k's to be continuous variables. In particular, some of the variables x_k can themselves be binary. For example, in liquefaction analysis, the jth explanatory variable might be defined as follows:

$$x_j = \begin{cases} 1 & \text{if the soil is silty sand} \\ 0 & \text{if the soil is clean sand} \end{cases} \tag{4.10}$$

Explanatory variables of this type are often called "indicator" variables, and they allow the consideration of qualitative attributes that are otherwise difficult to quantify, e.g. male/female, Democrat/Republican, or Japanese/non-Japanese earthquakes.

The third generalization involves the use of individual response data rather than grouped data. In the bio-assay example, the response of a group of insects, assumed to be homogeneous and administered the same "stimulus", was considered in the analysis. In most applications,

groups of homogeneous individuals are not available, and it is very difficult to obtain repetitions of the same "stimulus". This is certainly the case with the liquefaction data, where the "individuals" (sites) are heterogeneous and the "stimulus" (earthquake intensity) is not a variable that can be adjusted or controlled.

The coefficients β can be estimated directly from the individual responses and stimuli, without having to form groups of homogeneous individuals. The response Y of each individual is considered as a Bernoulli random variable with possible values Y=0 or Y=1. Such a scheme is easily incorporated into a maximum likelihood algorithm for estimating the β coefficients, the details of which are presented in Section 4.3.

So far, logit analysis has been discussed primarily in the context of a biological assay problem. However, analogs to the death/survival response of insects can easily be drawn to other situations. Relieved of the restrictions to the number and type (continuous/dichotomous) of the explanatory variables, and of having to form groups of homogeneous individuals or cases, a wide range of applications is possible. Examples of the use of logit regression in other disciplines are summarized in Table 4.1.

Within geotechnical engineering, the application of logit analysis has been reported in a paper by Veneziano and Liao (1984), which contains some of the preliminary results of the research on liquefaction presented in this thesis. At M.I.T., logit analysis has also been utilized in studies on slope stability (Carpenter, 1984), piping of soil particles (Honjo, 1985), and the use of neutron counts for the detection of rock joints (Einstein et al., 1985).

Most of the applications of logit analysis discussed in this section involve a binary response. An example of a problem involving a multinomial or polytymous response is the choice of transportation mode when several modes (auto, bus, train, etc.) are available; see Table 4.1. Techniques to deal with three or more forms of the response within the framework of logit or probit reqression can be found in numerous references, including Walker and Duncan (1967), Theil (1969), McFadden (1974), Daganzo (1979), and Ben-Akiva and Lerman (1985).

Liquefaction response could also be formulated in other than binary terms, e.g. following Tokimatsu and Yoshimi (1983) who suggest four classes of liquefaction behavior: no liquefaction, marginal liquefaction, moderate liquefaction, and severe liquefaction. Though the use of a polytomous response variable is an additional refinement, it also introduces more complex issues regarding the specification of the explanatory variables in modelling. For this reason, the study presented here is limited to the analysis of liquefaction as a binary response.

4.2.3. Logit versus Probit Analysis

As already noted, both the logit and probit transforms (Equations 4.3 and 4.4) are functions that "map" or convert values between zero and one into values between negative infinity ($-\infty$) and positive infinity ($+\infty$). There are other functions which can similarly be used to perform this mapping (see Cox, 1970, McFadden, 1976a, and Dobson, 1983), though they are generally less tractable or less theoretically justifiable for the analysis of categorical data. In the univariate case, the transforms basically convert symmetrical S-shaped curves into

straight lines, as illustrated previously in Fig. 1.1.

A comparison of the logistic function (used in logit analysis) and the normal or Gaussian cumulative distribution (used in probit analysis) is shown in Fig. 4.2. In this figure, both curves have been normalized to have zero mean and unit variance. It is evident that the functions are practically equivalent particularly in the central portions. Other properties of the logistic distribution can be found in Johnson and Kotz (1970).

The major argument for using logit analysis has always been the ease of computation and implementation. Note that the cumulative logistic distribution function has an explicit algebraic expression, whereas the normal cumulative function does not. Given also that the practical results are essentially the same, logit rather than probit was chosen as the method of analysis in this study.

4.2.4 Logistic Regression Versus Discriminant Analysis

In Chapter 2, the limitations of discriminant analysis in application to liquefaction risk assessement were discussed. This section presents a more detailed comparison of logit regression and discriminant analysis.

Logit analysis is a method of binary regression or curve fitting, whereas discriminant analysis is a method of classification. The two methods are related, in the sense that the results of logit regression can be used for classification, and the results of classification can be used (sometimes) to produce binary regressions. Both methods have appropriate uses, but in different settings and for different problems.

To clarify the situations for which binary regression and

discriminant analysis are each appropriate, consider the following set of examples (from McFadden,1976b):

- Example 1. (Causal Model): Seeds are planted, and observations are made on seed age, soil acidity, temperature, and time allowed for germination. Response: germination/no germination.
- Example 2. (Conjoint Model): Eggs are candled and observations are made on the translucency of the egg. Response: good egg (high yolk)/ bad egg (spread yolk).

McFadden (1976b) refers to the situation in Example 1 as a "causal model", i.e. the various factors observed are determinants of whether a seed germinates or not. Also important is the fact that it is not meaningful to speak of two seed populations, "germinators" and "non-germinators", that possess clearly defined distributions of control variables, i.e. seed age, soil acidity, etc. On the other hand, in Example 2, there are clearly two populations, "good eggs" and "bad eggs", which exist, but no causal relationships between translucency and whether an egg is "good" or "bad". Translucency and egg quality may be correlated, but they can be viewed as jointly determined by unobserved variables, and hence the nomenclature, "conjoint model".

Situations involving conjoint models are the proper candidates for discriminant analysis or more general classification techniques where the objective is to classify individuals or objects into one of two (or more) natural populations, e.g. good eggs/bad eggs, male/female, species 1/species 2. Problems involving causal models are more appropriately treated through binary regression. Liquefaction is

certainly typified by a causal model, where the intensity of earthquake shaking and the soil strength are independent variables that lead to the occurrence or non-occurrence of liquefaction. To analyze liquefaction data using classification methods is clearly incorrect.

The use of logit regression rather than discriminant analysis also overcomes a common objection by several researchers to the use of statistical methods to analyze liquefaction data. In trying to draw a boundary line separating liquefaction and non-liquefaction behavior, such as that shown in Fig. 1.1, it is argued that data points close to the line should be given more "weight". Data points far away from the line are considered to contain less information because they are not at critical locations that would control the position of the line. Physically, this notion corresponds to the fact that in many cases where a very severe earthquake shaking was experienced, the cyclic shear stresses at the site may have exceeded the soil liquefaction resistance many times over.

The above concern does present a problem within the framework of discriminant analysis, where two populations, "liquefiable sites" and "non-liquefiable" sites have to be assumed, and the underlying probability density functions of observed attributes are estimated based on the data. If there are "extremal" data points such as those where the liquefaction resistance was greatly exceeded, these data points would severely affect the estimates of the density functions and hence the location of the discriminant line. In the context of classification, the method of "least squares of misclassified points",

used by Yegian and Whitman (1978) was an attempt to address this problem. In logit analysis, the existence of "extremal" data points does not present a problem at all. In a sense, logit regression automatically gives more "weight" to those data points near the central portion (P = 0.5) of the logistic curve. Thus "extremal" data points affect the logit analysis to only a limited extent.

The preceding discussion of logit versus discriminant analysis has been on a methodological level. On a practical level, discriminant analysis simply does not produce the desired result of the conditional probability of liquefaction. Rather, it produces a decision rule to classify data into two groups, as illustrated by several examples in Chapter 2. Logistic regression, on the other hand, does produce the desired conditional probability.

4.3 Logit Analysis - Implementation

4.3.1 Maximum Likelihood Concepts

As stated previously, the objective of logit analysis is to use data observations to fit an equation of the form:

$$P = 1/[1 + \exp\{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m)\}] \qquad (4.11a)$$

or equivalently

$$\text{logit}(P) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m \qquad (4.11b)$$

where the x_k represent various explanatory variables and the β_k are regression coefficients to be obtained from analysis. Methods to obtain the β_k include a least-squares formulation which was originally

proposed by Berkson (1944). However the dominant trend is currently in favor of a maximum likelihood approach, in large part due to some desirable statistical properties of maximum likelihood estimates. These include:

- Consistency: As the number of data n becomes very large, the estimates of the coefficients β_k become unbiased and approach the correct values.
- Asymptotic Normality: As n becomes large, the estimates of the β_k become normally distributed.
- Asymptotic Efficiency: As n becomes large, the maximum likelihood estimator of β_k becomes efficient, i.e. it becomes the estimator with the minimum variance.

The last two properties are especially relevant in developing goodness-of-fit statistics for the logit models.

In maximum likelihood estimation, we assume a mathematical form (Eqn. 4.11) for the probability of occurrence of liquefaction, and then we ask: What combination of β_k's will give the maximum probability or likelihood of observing the actual pattern of liquefaction and non-liquefaction cases in the catalog? Solving this problem produces the maximum likelihood estimates of the β_k coefficients.

4.3.2 Likelihood Function

To mathematically formulate the likelihood function and the algorithm for finding the parameters that maximize the likelihood, the following modifications of Eqn. 4.11 are introduced. The probability of liquefaction for the ith observation (or case study) is denoted P_i with observed response i=1,...n, where n is the total number of

observations. Also, we need to distinguish between the observed explanatory variables x_k for each observation, and so we write x_{ik} which denotes the kth explanatory variable for the ith observation (i=1,...n). Thus,

$$P_i = 1/[1 + \exp\{-(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} + \ldots + \beta_m x_{im})\}] \quad (4.12)$$

is the probability of liquefaction for the ith case and $(1-P_i)$ is the corresponding probability of non-liquefaction. One can consider the response Y_i to be basically a Bernoulli random variable with probability P_i of the occurrence of $Y_i = 1$.

The likelihood or probability ℓ_i of observing either Y=1 or Y=0 for the ith case is simply:

$$\ell_i = P_i^{Y_i}(1-P_i)^{1-Y_i} \quad (4.13)$$

Note that since $Y_i = 0$ or $Y_i = 1$, $\ell_i = P_i$ in the case of liquefaction, and $\ell_i = (1-P_i)$ for non-liquefaction. If there are n independent observations, then the joint probability of occurrence of all observations is:

$$\ell = \ell_1 \cdot \ell_2 \cdot \ell_3 \ldots \ell_m = \prod_{i=1}^{m} \ell_i \quad (4.14a)$$

or

$$\ell = \prod_{i=1}^{m} P_i^{Y_i}(1-P_i)^{1-Y_i} \quad (4.14b)$$

This is the likelihood function that we seek to maximize with respect to the parameters β_k.

In practice, what is commonly done is to maximize the log of the

likelihood function rather than the likelihood function itself. This does not affect the values of the β estimates, since the logarithmic function is a monotonic one-to-one transformation. Hence we denote L as the log-likelihood function and write:

$$L = \ln(\ell) = \ln\left[\prod_{i=1}^{n} P_i^{Y_i}(1-P_i)^{1-Y_i}\right] \tag{4.15a}$$

or

$$L = \sum_{i=1}^{m} [Y_i \ln P_i + (1 - Y_i) \ln (1 - P_i)] \tag{4.15b}$$

Maximization entails taking the partial derivatives of L with respect to $\beta_k (k = 1, \ldots m)$ and setting them to zero. Also, we are interested in the 2nd derivatives of L as part of the Newton-Raphson algorithm of optimizing the log-likelihood and for deriving goodness-of-fit statistics. The first and second derivatives of L are easy to obtain and are given below:

$$\frac{\partial L}{\partial \beta_k} = \sum_{i=1}^{n} x_{ik}(Y_i - P_i) \qquad (k = 0, 1, \ldots m) \tag{4.16}$$

$$\frac{\partial^2 L}{\partial \beta_j \partial \beta_k} = \sum_{i=1}^{n} - x_{ij}x_{ik}(1 - P_i)P_i \qquad (j,k = 0, 1, \ldots m) \tag{4.17}$$

Note that $x_{i0} = 1$ for all i, and is associated with the intercept coefficient term β_0. In matrix notation, we can write the sets of these derivatives as:

$$\frac{\partial L}{\partial \beta_k} = \begin{bmatrix} \frac{\partial L}{\partial \beta_0} \\ \vdots \\ \frac{\partial L}{\partial \beta_m} \end{bmatrix} \tag{4.18}$$

and

$$\frac{\partial^2 L}{\partial \beta_j \partial \beta_k} = \begin{bmatrix} \frac{\partial^2 L}{\partial \beta_0^2} \cdots \frac{\partial^2 L}{\partial \beta_1 \partial \beta_m} \\ \vdots \\ \frac{\partial^2 L}{\partial \beta_m \partial x_o} \cdots \frac{\partial^2 L}{\partial^2 \beta_m} \end{bmatrix} = \underline{H} \tag{4.19}$$

The matrix of 2nd order partial derivatives (Eqn. 4.19) is sometimes referred to as the Hessian matrix, and denoted $\underline{H}$.

4.3.3. Likelihood Maximization and Computer Codes

To obtain the maximum likelihood estimates of β_k (where $\partial L/\partial \beta_k = 0$, for all k), an iterative solution based on the Newton-Raphson formula is used. At the rth step, one calculates:

$$\{\beta_k\}_{r+1} = \{\beta_k\}_r - \left[\frac{\partial^2 L}{\partial \beta_j \cdot \partial \beta_k}\right]_r^{-1} \left\{\frac{\partial L}{\partial \beta_k}\right\}_r \tag{4.20}$$

At the zero-th step, $\{\beta_k\}$ is usually initialized by setting $\beta_k = 0$ for all k. Convergence is achieved usually within 5 to 10 iterations.

Methods other than the Newton-Raphson algorithm can be used to obtain the maximum likelihood solution. These methods are often implemented for problems where explicit closed-form expressions for the

derivatives of the log-likelihood functions are difficult to obtain; one example is the general multinomial probit model (Daganzo, 1979). For the binary logit model, the Newton-Raphson algorithm is computationally efficient and is easy to implement.

There are several logit regression programs that are generally available, usually as part of a larger package of statistical programs. QUAIL was used initially in the preliminary studies of liquefaction data presented by Veneziano and Liao (1984). However, the input and command structure required for QUAIL was far too complex and cumbersome for the application to simple binary logistic regression. [QUAIL is intended primarily for use with multinomial response data.] In addition, part of the research required modifications of the standard logit regression procedure (see Chapter 6) and it was deemed preferable to write a new computer code than to change a pre-existing one. A FORTRAN listing of the logit regression program written for this study is given in Appendix B.

4.3.4 Goodness-of-Fit Statistics

Goodness-of-fit statistics for logit regression are useful for two purposes: 1) to decide whether a proposed model is statistically significant and 2) to compare various competing models in trying to decide which is "best". By the term "model", we mean the particular combination of explanatory variables (e.g. cyclic stress ratio and SPT resistance) that constitute the x_k's in the logistic formula (Eqn. 4.11). The term "goodness of fit statistics" is used here to mean a variety of statistics in model fitting and parameter estimation, though <u>not all</u> are goodness-of-fit statistics in the strict sense of the term.

Several statistics are available and are briefly reviewed here. Of these, the modified likelihood ratio index (denoted by MLRI or $\bar{\rho}^2$) was found to be the most useful in the analysis of liquefaction data.

Percent Correctly Predicted (PCP): A response for the ith data point is said to be correctly predicted if $Y_i = 1$ and $P_i > 0.5$; or if $Y_i = 0$ and $P_i < 0.5$. P_i is evaluated using Eqn. 4.11 with the values of β_k which maximize likelihood function. The PCP statistic has an intuitive meaning but has been shown by Horowitz (1982) to be neither a very discriminating nor a particularly useful statistic in model comparison. This is because PCP depends only on the 0.5 probability value and not on the entire logistic probability function.

T-Statistics for the β Parameters: Let $\hat{\beta}_k$ denote the maximum-likelihood estimate of β_k and $\hat{\sigma}_k$ denote the standard deviation estimate of $\hat{\beta}_k$. Then the null hypothesis that $\beta_k = 0$ can be tested by comparing:

$$T = \frac{|\hat{\beta}_k|}{\hat{\sigma}_k} \tag{4.21}$$

with the $(1-\alpha/2)$ fractile of the Student's t distribution with $n-(m+1)$ degrees of freedom, where n is the total number of observations (of liquefaction and non-liquefaction) and m is the number of explanatory variables. For large $n-(m+1)$, the T-statistic can be assumed to be approximately normal, which yields the result that the null hypothesis is significant at the α level (2-sided test) for $T > \Phi_{1-\alpha/2}$ (e.g. for $\alpha = 0.05$, $T > 1.95$). This implies that for β-coefficients with values of T greater than or equal to about 2.0, these parameters are statistically significant at the 0.05 level.

The standard deviations $\hat{\sigma}_k$ are obtained from the diagonal elements of $-\underline{H}^{-1}$, where $\underline{H}$ is the Hessian matrix (Eqn. 4.19) evaluated at the maximum likelihood β_k parameters. This result is a consequence of the Cramer-Rao inequality (see Theil, 1971).

Likelihood Ratio Statistics LR_0 and LR_c: One can define several likelihood ratio statistics. The ones presented here are used to test the significance of the maximum likelihood model against two "naive" models (null hypotheses). LR_0 is used to test the maximum likelihood model against the model where all the coefficients $\beta_k = 0$, i.e. against $P \equiv 0.5$ for all observations. The statistic LR_c is used to test the fitted model against the null hypothesis that the probability of Y=1 (liquefaction) is simply equal to P_c, the observed fraction of liquefaction cases. The two statistics are defined as:

$$LR_0 = -2 \ln \left[\frac{\ell(0)}{\ell(\hat{\beta})} \right] = -2 \left[L(0) - L(\hat{\beta}) \right] \tag{4.22}$$

$$LR_c = -2 \ln \left[\frac{\ell(c)}{\ell(\hat{\beta})} \right] = -2 \left[L(c) - L(\hat{\beta}) \right] \tag{4.23}$$

where $L(\hat{\beta})$ is the log likelihood evaluated using the fitted parameters $\hat{\beta}_k$; $L(0)$ is the log likelihood evaluated assuming $\beta_k = 0$; $L(c)$ is the log likelihood evaluted assuming a single constant parameter $c = \ln[P_c/(1-P_c)]$.

The statistics LR_0 and LR_c are both asymptotically χ^2 distributed with respectively (m+1) and m degrees of freedom, where m is the number of explanatory variables. Simple formulas can be derived for and L(c) and are given below (for binary logit):

$$L(0) = -n \ln(2) \tag{4.24}$$

$$L(c) = \ln \left[\frac{n_1^{n_1} n_o^{n_o}}{n^n} \right] = n_1 \ln n_1 + n_2 \ln n_2 - n \ln n \qquad (4.25)$$

where n_1 is the number of cases for which Y = 1; n_o is the number of cases for which Y = 0; and $n = n_o + n_1$ is the total number of observations.

The hypotheses which are tested using LR_o and LR_c are that the maximum likelihood fits are no better than the respective naive models. The models are significant at level α against the "naive hypotheses" if $LR_\alpha > \chi^2_{\alpha,m+1}$ or $LR_c > \chi^2_{\alpha,m}$ (where the symbol $\chi^2_{\alpha,\nu}$ denotes the $(1-\alpha)$ fractile of the χ^2 distribution with ν degrees of freedom. In this study, the significant level used was $\alpha = 0.05$.

Modified Likelihood Ratio Index (MLRI or $\bar{\rho}^2$): This statistic was proposed by Horowitz (1982) and is defined for binary logit as:

$$\bar{\rho}^2 = 1 - \frac{L(\beta) - (m+1)/2}{L(0)} \qquad (4.26)$$

where $L(\beta)$ denotes the log-likelihood function evaluated using the maximum likelihood values of β_k. $L(0)$ denotes the log-likelihood function assuming $\beta_k = 0$ (for all k); and m is the number of explanatory variables (Note: m+1 is the number of β coefficients or "parameters", i.e. m variables plus the intercept β_o). The MLRI or $\bar{\rho}^2$ statistic is based on the "unmodified" likelihood ratio index (LRI or ρ^2) defined in a similar way (McFadden, 1974), but without the term $(m+1)/2$, which represents a correction to account for the number of

parameters in the model.

The modified likelihood ratio index was devised to compare "non-nested" models, i.e. models such that one model cannot be obtained from the other by eliminating one or more explanatory variables. "Nested" models can be compared using various likelihood ratio statistics (see Sec. 4.3.5). The values of $\bar{\rho}^2$ vary theoretically between 0 and 1, with higher values indicating better fit to the data. This is analogous to the correlation coefficient R^2 in linear regression, but whereas "good" fits in linear regression correspond to values of R^2 of at least 0.8, a good logistic regression fit corresponds to much lower values of $\bar{\rho}^2$. Within the applications of transportation and consumer demand analysis, values of $\bar{\rho}^2$ on the order of 0.3 to 0.4 are considered to "excellent" (Hensher and Johnson, 1981; Lerman, 1983).

Horowitz (1983) has studied the significance that can be attributed to a difference in $\bar{\rho}^2$ between 2 models. Let $\bar{\rho}_A^2$ and $\bar{\rho}_B^2$ denote the modified likelihood ratio indices for a 2 non-nested models A and B, and suppose that $\bar{\rho}_A^2 > \bar{\rho}_B^2$. The probability that B is actually the correct model (even though it has a larger $\bar{\rho}^2$) can be estimated by the upper bound result:

$$\Pr(\bar{\rho}_A^2 - \bar{\rho}_B^2 > \Delta\bar{\rho}^2) \geq \Phi[-\{2n(\Delta\bar{\rho}^2)\cdot\ln(2)\}^{1/2}] \qquad (4.27)$$

where $\Delta\bar{\rho}^2$ is an arbitrary value (and can be set to the calculated difference in $\bar{\rho}^2$), and Φ denotes the the standard normal cumulative distribution function.

Based on Horowitz's result, the upper bound values of $\Delta\bar{\rho}^2$ between

two models is plotted in Fig. 4.3. For a signficance level $\alpha = 0.05$ and 200 observations, $\Delta\bar{\rho}^2$ is less than 0.01. This means that for 200 observations, if 2 models are being compared and one of them has a $\bar{\rho}^2$ greater by 0.01, the model with the higher $\bar{\rho}^2$ is superior at a significance level of 0.05. This illustrates that $\bar{\rho}^2$ is an excellent and discriminating statistic for comparing different models.

4.3.5 Stepwise Logistic Regression

In logit analysis, one is often faced with a large number of possible models, i.e. of possible combinations of explanatory variables. From the various candidate models, we would like to select one or a few that are optimal in some sense. To check all the possible combinations of explanatory variables is a tedious task, but the search can be automated to a certain degree by a stepwise regression procedure, e.g. as described in Anderson (1982) and Feinberg (1980).

The "stepwise" portion of the logistic regression program used in this study was coded by Honjo (1985), and his implementation is described below. A set of candidate variables is hypothesized, and an initial model (i.e. an initial subset of variables) is chosen. The significance level α of the desired final model must also be specified. The program then tries to build up (forward steps) or to pare down (backward steps) the model to obtain a statistically optimum model with only a few explanatory variables. At each step, the procedure either adds the most statistically signficant variable not already included in the "current" model, and/or deletes the least significant variable from the "current" model.

The following likelihood ratio criteria is used to determine

whether a variable is statistically significant or insignificant:

$$LR = -2 \ln\left(\frac{\ell_{m+1}}{\ell_m}\right) = -2(L_m - L_{m+1}) \tag{4.28}$$

In Eqn. 4.28, ℓ and L respectively denote the likelihood and log-likelihood functions and the subscripts m and m+1 denote the number of explanatory variables in the two "nested" models under consideration. LR is χ^2 distributed with one degree of freedom. If $LR > \chi^2_{\alpha,1}$, then the candidate variable is added in a forward step or exempted from deletion in a backward step.

The choice of the significance level α is somewhat arbitrary, and is based on judgment. In this study, after some initial trials, an $\alpha = 0.05$ was selected because, at this level, the stepwise rgression procedure yielded results that were considered to be consistent with physical considerations.

Two important points are mentioned here with respect to the interpretation of the results of stepwise analysis. First, the optimization procedure is local, not global, meaning that convergence may be to a locally optimum model. Good practice is therefore to obtain several trial results by working with different initial subsets of variables. Second, the optimization procedure may yield models that are statistically significant, and yet contain terms that are not relevant from a physical viewpoint. Thus, the results of statistical analysis may need to be further scrutinized and possibly corrected based on physical understanding of the liquefaction phenomenon.

4.4 Logistic Models of Liquefaction Behavior

4.4.1 Model Formulation and Interpretation

Although the present approach to liquefaction risk is statistical, it is desirable that the form of the model (fitted to the data) agree, at least qualitatively, with the current physical understanding of the phenomenon. In particular, it may happen that a variable which does not have an evident physical interpretation is found to be statistically significant. In formulating and interpreting logistic liquefaction models, consideration was given to the following issues:

1) Statistical significance.
2) Accuracy of estimation (i.e. are there enough data to reliably identify the structure of the model and to evaluate the results?)
3) Physical interpretation.
4) Biases due to particular features of the data.

The first two issues are easy to address through hypothesis testing and the calculation of statistics of the type presented in Section 4.3.4. Issues (3) and (4) are somewhat more complicated and require a physical understanding of the liquefaction phenomenon. Judgement is necessary to select physically significant models from among those found to be statistically significant.

4.4.2 Base Local and Global Models

A goal of the research was to concurrently develop two types of statistical models of liquefaction behavior -- one based on "local" and the other based on "global" earthquake intensity measures. As

discussed previously in Chapter 2, one may distinguish between liquefaction models that use the local site intensity measures (e.g. peak acceleration and duration) and models that use a global source characterization (e.g. magnitude and epicentral distance). These are sometimes also referred to respectively as A&D and M&R models (Yegian and Whitman, 1978). The physical models selected to form the basis of these two types of the logistic models are referred to as the Seed-Idriss and the Davis-Berrill "base models". The use of the terminology "base model" is meant to distinguish it from the actual physical model. Also, the "base models" are fairly simple, but they will form the basis of more complex models to be discussed subsequently. In instances where the context is clear, the term "base" will sometimes be omitted for the sake of brevity.

The selected base local (A&D) model corresponds to the method of liquefaction analysis originally proposed by Seed and Idriss (1971). Recent refinements of their method include an earthquake magnitude normalization factor for the cyclic stress ratio and a normalization of SPT resistance to a standard hammer energy (Seed, et al., 1984). The Seed-Idriss model is semi-empirical, being based in part on analytical results, and in part on empirical measurements. The Seed-Idriss model was chosen as a base model because it is the method of liquefaction analysis most often used in practice, particularly in the United States.

The selected base global (M&R) model incorporates an earthquake load parameter originally proposed by Davis and Berrill (1981,1982). This model is also semi-empirical and is formulated on the basis of energy

dissipation considerations. In contrast to the Seed-Idriss model, there were no a priori reasons, such as prevalent usage by the profession, to warrant its selection as the base M&R model. Rather, the Davis-Berrill model was chosen from among several candidate M&R models largely because of its superior goodness-of-fit to the data. The details of this selection process are documented in Section 4.4.4.

Both the Seed-Idriss and the Davis-Berrill models reduce the liquefaction criterion to the comparison of two variables, one representing earthquake load and the other a measure of the soil liquefaction resistance. In both models, the resistance functions are the SPT N-value, normalized or corrected for an overburden pressure of 1 kg/cm^2 (which is approximately equal to 1 TSF or 100 kPa). The corrected N-value is denoted by N_1, and the N_1 value adjusted to reflect a standard 60% hammer energy efficiency is denoted $(N_1)_{60}$, as per Seed et al. (1984). [See Chapter 3 for details.]

The load variable in the Seed-Idriss model is the normalized cyclic stress ratio, CSRN defined as:

$$CSRN = \left(\frac{\tau}{\bar{\sigma}_v}\right)_{M=7.5} = 0.65 \frac{a}{g} \frac{\sigma_v}{\bar{\sigma}_v} \frac{r_d}{r_m} \qquad (4.29)$$

where a is peak surface acceleration, g is the acceleration constant σ_v and $\bar{\sigma}_v$ are respectively the total and effective overburden stresses, r_d is a depth reduction factor and r_m is a normalization factor to convert the cyclic stress ratio (CSR) for an arbitrary magnitude M to CSR for an equivalent M = 7.5 earthquake (see Seed et al., 1984, or Chapter 3). CSRN is a dimensionless quantity and arbitrary units can be used for

a, σ_v and $\bar{\sigma}_v$.

The load function Λ for the Davis-Berrill model is defined as:

$$\Lambda = \frac{10^{1.5M}}{R^2(\bar{\sigma}_v)^{3/2}} \tag{4.30}$$

where M is the earthquake magnitude, R is a distance measured from the source to the site, and $\bar{\sigma}_v$ is the effective overburden stress. Note that Λ in Eqn. 4.30 is actually the inverse of the function used by Davis and Berrill (1982), and that Λ is not dimensionless. Therefore, the numerical value of Λ depends on the units of R and $\bar{\sigma}_v$, which are taken here to be kilometers and kg/cm^2 respectively. In some cases, it is convenient to operate on a dimensionless quantity λ which is defined as:

$$\lambda = \Lambda/\Lambda_0 \tag{4.31}$$

where Λ_0 is Λ evaluated for M = 7.5, R = 100 km and $\bar{\sigma}_v$ = 1 kg/cm^2.

The base logistic regression models that correspond to the physical models are for Seed-Idriss:

$$\text{logit}(P) = \beta_0 + \beta_1 \ln(CSRN) + \beta_2(N_1)_{60} \tag{4.32a}$$

or

$$P = 1/[1 + \exp\{-(\beta_0 + \beta_1 \ln(CSRN) + \beta_2(N_1)_{60})\}] \tag{4.32b}$$

and for Davis-Berrill:

$$\text{logit}(P) = \beta_0 + \beta_1 \ln \Lambda + \beta_2(N_1)_{60} \tag{4.33a}$$

or

$$P = 1/[1 + \exp\{-(\beta_0 + \beta_1 \ln\Lambda + \beta_2(N_1)_{60})\}] \tag{4.33b}$$

Reasons for these particular functions and for the choice of the

Davis-Berrill model over other M&R candidate models are discussed in the following sections.

4.4.3 Justification of the Parametric Form

The logit regression models described in this chapter are all of the "linear" type because in the logit scale, the regression equation (Eqn. 4.3 or 4.11b) is linear with respect to the explanatory variables x_k. The base logit models for the Seed-Idriss and Davis-Berrill formulations are both of the form:

$$\text{logit}(P) = \beta_o + \beta_1 x_1 + \beta_2 x_2 \qquad (4.34)$$

where $x_1 = \ln(\text{CSRN})$ or $x_1 = \ln \Lambda$ and $x_2 = (N_1)_{60}$. Notice that although logit(P) is linear in x_1 and x_2, the x_1 term is actually a non-linear function of explanatory load variable CSRN or Λ. However, the model is still a linear logistic equation. In the proposed formulation, CSRN and Λ are essentially transformed by the natural log function to a different scale.

The question naturally arises as to which functions of the physical variables should be used in the model. In practice, the choice is often made by trial and error guided by considerations of both physical meaning and statistical significance. In Table 4.2 and Fig. 4.4, the goodness-of-fit statistics and implications of various functional forms considered in formulating the Seed-Idriss based model are compared. In particular, it is clear from Fig. 4.4 that the use of different functional forms for x_k may have important consequences on liquefaction risk predictions. For example, if one defines $x_1 = \text{CSRN}$ and $x_2 = (N_1)_{60}$, then the model gives constant probability of liquefaction along parallel

straight lines on the CSRN vs $(N_1)_{60}$ plane as shown in Fig. 4.4(a). Other definitions of x_1 and x_2 yield contour lines of P with different shapes. The contour lines, in particular the 0.5 contour lines, can be compared to classification lines derived from discriminant analysis.

Based on generally accepted notions of the physical mechanism of liquefaction, the results shown in Fig. 4.4(a) and 4.4(b) are clearly inappropriate. More reasonable models are those represented in Figures 4.4(c) and 4.4(d). In particular, the model of Fig. 4.4(c) corresponds to Eqn. 4.32 and is considered to be the most appropriate. The associated equiprobability lines have the equations of the form:

$$CSRN = b \cdot \exp\{c \cdot (N_1)_{60}\} \tag{4.35}$$

where b and c are constants that depend on the probability of liquefaction. In the case of the model shown in Fig. 4.4(d), the contour lines have equations of the type:

$$CSRN = b \cdot [(N_1)_{60}]^c \tag{4.36}$$

and are obtained by fitting the logistic model:

$$\text{logit}(P) = \beta_o + \beta_1 \ln(CSRN) + \beta_2 \ln[(N_1)_{60}] \tag{4.37}$$

The difference between Eqns. 4.32 and Eqn. 4.37 is in the defintions of x_2 as either $(N_1)_{60}$ or $\ln[(N_1)_{60}]$; these equations will be referred to in the discussion that follows as the $(N_1)_{60}$ and $\ln[(N_1)_{60}]$ models, respectively.

The $(N_1)_{60}$ model (Fig. 4.4(c)) produces a concave upward shape for the equiprobability lines, whereas the $\ln[(N_1)_{60}]$ model (Fig. 4.4(d))

gives rise to a concave downward shape. That the curvature of the contour lines is a contentious issue was illustrated earlier in Chapter 1 (Fig. 1.2). Though a concave upward shape is the generally accepted notion, there are several exceptions. For example, Tokimatsu and Yoshimi (1981) have noted that the method proposed by Iwasaki, Tatsuoka, et al. (1978) implies a concave downward curvature. The result obtained by Yegian and Vitelli (1981a, 1981b) also has the same implication.

Another difference between the two models is that the equi-probability contour lines pass through the origin in the $\ln[(N_1)_{60}]$ model whereas they have a positive CSRN-axis intercept in the $(N_1)_{60}$ model. This convergence of the contour lines in the $\ln[(N_1)_{60}]$ model is a natural constraint imposed by the model, and this constraint naturally leads to the concave downward curvature. The exponent c in Eqn. 4.36, which determines the curvature of the contour lines equals the ratio β_2/β_1 of the logit coefficients in Eqn. 4.27. The ratio found from fitting to the data is smaller than 1.0, resulting in the concave downward curvature. If the data allowed, the exponent could have well been greater than 1.0 and concave upward contours would have been the result.

Consider the schematic shown in Fig. 4.5 for the case of a polynomial curve-fitting problem. In Fig. 4.5(a), the data points and a first order polynomial (straight line) is shown to fit the data fairly well, but with a positive y-axis intercept. If the same data are then fitted to a higher order polynomial (or a power function) constrained to go through the origin as shown in Fig. 4.5(b), a concave downward line results. In this schematic, it is clear that the concave downward

curvature of the polynomial is an artifact of the constraint that the polynomial line pass through the origin. It is conjectured that this same sort of artifact affects the $\ln[(N_1)_{60}]$ logit model, and also possibly the results obtained by others showing a concave curvature of lines separating liquefaction and non-liquefaction in the CSRN-N_1 space.

The prediction of the proposed $(N_1)_{60}$ model (Fig. 4.4(c)) is that a threshold value of CSRN is required to cause liquefaction even at very low values of soil strength. The non-zero CSRN, axis intercept of the equiprobability lines presents somewhat of a dilemma. On the one hand, there are several precedents that support such a phenomenon: Seed et al. (1984) indicate separating lines trending toward a non-zero intercept -- at least for silty soils. Bierswale and Stokoe (1984) has shown liquefaction criteria on a plot of acceleration versus soil shear wave velocity in which the separating lines trend towards a threshold acceleration; the M&R model proposed by Yegian and Whitman (1978) also implies a similar intercept. On the other hand, it could be argued that the contour lines should go through the origin. Physical intuition leads us to envision that at $(N_1)_{60} \approx 0$, i.e. almost zero soil strength, there would be a correspondingly small cyclic shear stress ratio CSRN required to cause liquefaction.

There are three possible explanations for the appearance of a trend toward a non-zero CSRN intercept of separating or equiprobability lines. The first is that there might actually be a change from concave upward to downward curvature which could insure that the lines would pass through the origin. In this case, models fitted for higher soil resistance data should not be extrapolated back to very low resistances. The second

possibility is that the intercept is indeed non-zero as explained by the threshold strain concept proposed by Dobry et al. (1982). According to these authors, there is a critical level of cyclic strain of about 10^{-2}% or 0.0001, below which no excess pore water pressure buildup is possible, thus rendering liquefaction nearly impossible. Their calculations indicate that even for sites with soils of very low strength, a threshold acceleration level of about 0.05g is required to cause liquefaction. The third possible explanation of the non-zero CSRN intercept is based on the way the SPT is performed and the meaning of a N-value equal to zero, as discussed below.

In performing the test, the borehole is advanced to the desired depth, the SPT sampler is lowered into the hole, and the combined weight of the drilling rods and the 140 lb SPT hammer (seated on the anvil) is statically brought to bear on the soil at the tip of the sampler. Then the hammer is repeatedly raised a standard height of 30 inches and dropped until the sampler has advanced 18 inches into the soil. The number of hammer blows required to drive the sampler the last 12 inches is recorded as N, with the blows attributed to the first 6 inches noted as "seating" blows. To obtain an N-value of 1, the sampler has to be first "seated" 6 inches, and then one drop of the hammer should advance the sampler another 12 inches. To obtain N = 1/2, the hammer drop would have to advance the sampler by 24 inches, or one could drop the hammer from only 1/2 its standard height to advance it 12 inches. To get close to a measurement of N = 0, one would have to drop a hammer from an extremely small height, create an impact, and advance the sampler exactly 12 inches.

The point is that even a value of N ≈ 0 does not indicate that the soil being tested has zero strength. Rather, the soil must at least be able to withstand the static weight of the rods and the hammer, and the sampler has to be "seated" into the soil, before N can be measured. It is emphasized that the notation WOR (Weight of Rods) or WOH (Weight of Hammer) often used on boring logs to indicate that the soil could not even withstand the static weight of the rods and/or the hammer is a better indication of zero soil strength, but this does not correspond to N = 0.

In the opinion of the writer, the threshold strain concept combined with the fact that N = 0 does not mean zero soil strength, give convincing arguments in favor of a non-zero CSRN intercept for the equiprobability lines. It is also reassuring to see from Table 4.2 that the MLRI or $\bar{\rho}^2$ goodness-of-fit indicator also favors the use of x_1 = ln(CSRN) and $x_2 = (N_1)_{60}$ as explanatory variables.

The equi-probability contours shown in Fig. 4.4 have been obtained by fitting the model to all the 278 cases from the catalog without distinction between silty sands and clean sands. At present, the type of sand is considered to be a factor with important effects on liquefaction resistance (Seed et al., 1984). The above-noted difference in the curvature of the equi-probability lines obtained from different models were also noted when the data were divided into clean sand and sitly sand subsets. This is shown in Fig. 4.6. A sand is considered "clean" if it has fines content FC < 12%, and "silty" if FC > 12%. The justification for this classification criterion is given in Section 4.8.1.

A similar series of candidate functional forms were also fitted to

the Davis-Berrill model (i.e., considering $x_1 = \ln\Lambda$ and either $x_2 = (N_1)_{60}$ or $x_2 = \ln[(N_1)_{60}]$). The physical implications of the shapes of the equi-probability were less compelling in indicating which model was preferable, because it is less clear from physical reasoning what shapes these curves should have. On a statistical basis, the MLRI statistic indicates that the base model proposed in Eqn. 4.33 be the one that fits the data best. Using $x_2 = (N_1)_{60}$ in both the Davis-Berrill and Seed-Idriss base models also makes the two formulations consistent.

4.4.4 Davis-Berill versus Other M&R Models

This section documents the reasons behind choosing the Davis-Berrill formulation as the base M&R model used in this study. Several candidate models were considered, including those proposed by Yegian and Whitman (1978) and Yegian and Vitelli (1981a, 1981b) shown in Table 4.4. A more general M&R model based almost entirely on the data (i.e. with minimal physical considerations) was also considered.

The most general M&R logit model used in the analysis can be represented by the regression equation:

$$\text{logit}(P) = \beta_o + b_1 M - b_2 \ln R + b_3 \ln\sigma_v - b_4 \ln\bar{\sigma}_v + \beta_2 (N_1)_{60} \qquad (4.38)$$

where the b_k's are positive coefficient and R is a suitably defined measure of distance between the earthquake source and the site of interest, e.g. epicentral or hypocentral distance. This equation can be rewritten as:

$$\text{logit}(P) = \beta_o + \ln \Psi + \beta_2 (N_1)_{60} \qquad (4.39)$$

where

$$\Psi = \frac{e^{b_1 M}}{R^{b_2}} \frac{\sigma_v^{b_3}}{\bar{\sigma}_v^{b_4}} \qquad (4.40)$$

Therefore, Eqn 4.38 is a generalization of the proposed Davis-Berrill base model (Eqn. 4.33) and Ψ (in Eqn. 4.40) is a generalized form of the parameter Λ in that model as well as any of those shown in Table 4.4. The power coefficient b_3 for the total stress σ_v was indicated by analysis to be statistically insignificant and to be justifiably set equal to zero. There is also an implied coefficient $\beta_1 = 1$ in Eqn. 4.39 for the ln Ψ term, which however does not diminish the generality of the model.

The essential difference between using the generalized model and of the other M&R models in Table 4.4 is that the b-coefficients (in Eqn. 4.38 or Eqn. 4.40) are fixed in the latter cases. In the generalized model, the b-coefficients are to be estimated from the data in the same way as the β-coefficients. Thus, it is naturally expected that any model with pre-assigned b-coefficients can be no better than the generalized model.

The results from fitting various logit regression models of the global M&R type are shown in Table 4.5. The model showing the best statistical fit is the generalized model with R defined as the hypocentral distance ($R = R_{HY}$). The modified Davis-Berrill model with $R = R_{HY}$ also has comparably good statistics, whereas the Yegian and Whitman (1978) and Yegian and Vitelli (1981b) models produce poorer fits to the data. The logit regression results are presented in Table 4.6 for

comparison in a slightly different format, where generalized load function Ψ is written for each of the same models as in Table 4.5. Comparing the power coefficients in each of the equations leads to the conclusions that the Davis-Berrill and the generalized models are in good agreement, whereas the power coefficients of the Yegian-Whitman and Yegian-Vitelli models, are markedly different (particularly in the coefficient that accounts for earthquake magnitude).

Numerous other candidate load functions were tried, primarily involving variations in the definition of the distance measure R. These models were found to be statistically suboptimal and the results are not reported here. One particular set of models with R defined as the distance to energy release (DER) is worth mentioning because it is a recurrent issue in liquefaction analysis using M&R models (Youd, 1977; Davis and Berrill, 1982). In all cases where the distance to energy release (DER) was used in place of epicentral distance R_{EP} or hypocentral distance R_{HY}, the DER formulation was found to fit the data relatively poorly, with the MLRI statisic lower by 0.1 or more, when compared with either the R_{EP} or R_{HY} models.

Recent studies (Joyner and Boore, 1981; Campbell, 1981) indicate that the DER measure is an appropriate choice in the fitting of attentuation relationships to strong motion data. It is therefore surprising that DER formulations do not perform as well as R_{EP} and R_{HY} formulations. It is conjectured that this is due to inaccuracies in determining DER particularly in regions where surface expressions of the fault rupture are not evident. Future studies using more accurate evaluations of DER may yield a different conclusion. However, at present

there is no indication that DER is a good choice as a distance measure in M&R liquefaction models.

Another point of interest is the comparison of the M&R models and the Seed-Idriss (A&D) model. The MLRI statistic for the Seed-Idriss model fitted to the same data subset as that used for the results in Table 4.5 was 0.5013 compared to 0.4677 for the generalized M&R model with $R=R_{HY}$. This difference indicates the Seed-Idriss model is significantly better than the statistically "best" M&R model when local accelerations can be estimated accurately. However, this is not a justification for using the Seed-Idriss model in risk analysis, especially if peak site acceleration has to be determined from an attentuation relationship. Uncertainties in the attenuation relationship will invariably affect estimates of local accelerations, and M&R models may perform better in such cases.

The reasons for choosing the Davis-Berrill model as the basis for more complex formulations in logit analysis are:

1) The Davis-Berrill Model (whether one uses the epicentral or the hypocentral formulation) fits the data as well as a more elaborate M&R model.
2) The Davis-Berrill Model is based on a plausible (though approximate) physical model of liquefaction behavior and thus provides a physical interpretation of the results of logit analysis.

Study of both the hypocentral and epicentral models are pursued further in later sections of this thesis. The hypocentral (Modified) Davis-Berrill model is a simple extension and and actually fits the data somewhat better than the model based on the earthquake load parameter

originally proposed by Davis and Berrill (1981, 1982). However, it may not be directly applicable for liquefaction risk analyses in regions where seismicity is poorly understood so as to make estimates of earthquake focal depths highly uncertain. In such cases, use of a model using the original Davis-Berrill load parameter which relies on an epicentral distance measure would be more appropriate.

Contour plots of probability of liquefaction are shown in Fig. 4.7 for the epicentral and hypocentral (Modified) Davis-Berrill models. In this figure, the contours are straight lines because of the logarithmic scale chosen for the load parameters Λ_{EP} and Λ_{HY}. The expression for describing the probability of liquefaction for the epicentral Davis-Berrill model is:

$$\text{logit}(P) = -12.922 + 0.87213 \ln(\Lambda_{EP}) - 0.21056 (N_1)_{60} \qquad (4.41)$$

which is based on a regression using all 278 case study data. The corresponding equation for the Modified Davis-Berrill model (based on 228 cases with available focal depth data) is reported in Table 4.5.

4.5 Sensitivity of the Results to the Data Sets

A limitation of the liquefaction catalog discussed in Chapter 3 is that not all the case studies provide a complete set of data. For example, of the 278 total cases, only 159 have an indication of the soil fines content (FC) and even fewer (113) have a complete set of all the soil gradation characteristics compiled for this study. Thus, in extending the base Seed-Idriss model to include the effects of fines content (FC), the subset of 159 data which has reported fines content

has been employed. On the other hand, if one wanted a model incorporating the effect of all gradation characteristics, only 113 case studies would be available.

In comparing the results obtained using different models based on different subsets of data, the question arises as to whether the results are consistent. If one partitions the data to create subsets in a way that is not biased, then the results obtained from the various subsets should be comparable. This issue is similar to a concern regarding different conclusions that might be derived by different investigators who have each compiled their own liquefaction data catalogs.

Fig. 4.8 shows a comparison of the logit regression coefficients β_k obtained for the Seed-Idriss model using various subsets of the data. A regression model fitted to a recent data catalog compiled by Seed et al. (1984) is also shown. Reasonably narrow ranges of estimates of the β_k are indicated and are comparable to the one standard deviation error bars shown for the individual estimates. The lines of equal probability of liquefaction in the CSRN-$(N_1)_{60}$ plane that corresponding to the different coefficient estimates are shown in Fig. 4.9. In this figure the contours are seen to be fairly similar regardless of the data set or subset that was used in the analysis. A superimposed comparison of the 0.1, 0.5, and 0.9 equi-probability contours for each of the data sets/subsets is also shown in Fig. 4.10.

Figures 4.8 through 4.10 show results of the Seed-Idriss base model fitted to the data sets without the consideration of the fines content of the soil. In Figures 4.11 through 4.14, similar results are shown for

comparable data partitions, but with the data further divided into subsets of clean sands with fines content FC < 12% and silty sands with FC > 12%. The rationale for such a dichotomy into clean and silty sands is discussed in Section 4.8.1. Using the Davis-Berrill base model, similarly small ranges of β_k and contour lines were also found when the model was fitted to various data sets.

Uncertainty on the location of the equi-probability lines can be couched in the more formal context of deriving confidence intervals. There is an exact solution for the case of analysis involving one explanatory variable (Brand et al, 1973), and an approximate solution for more than one explanatory variable has been proposed by Hauck (1983). Honjo (1984) investigated the use of Hauck's method to obtain confidence bands, but found it to be inaccurate, except possibly for probability values near 0.5. Thus at present, there are no reliable methods for obtaining confidence bands for general multivariate binary logit. Alternative ways of obtaining confidence bands for logit analysis would be to use statistical procedures such as "jackknife" and "bootstrap" methods (see Efron, 1982, or Efron and Gong, 1983). The idea behind these methods is to create artificial data sets by resampling from the original data, refit the model to each data set and characterize uncertainty of the parameters (in our case the logit coefficients) on the basis of their variability from set to set. The procedure described earlier in comparing logit regression results obtained from different data sets is similar in concept though not in execution. It should also be noted that bootstrap and jackknife methods usually involve a considerable computational effort.

In summary, the results presented in this section give an idea of the variability of liquefaction probability estimates that one obtained by using various data sets. Based on the results and considering the accuracy needed for making engineering decisions, it appears that the variation is not too large.

4.6 Extensions of the Base Logit Models

The Seed-Idriss and Davis-Berrill base models provide rather good fits to the data. However, there are also several additional explanatory variables that one might wish to consider. Examples of such variables are soil gradation characteristics, the method of estimating accelerations at a site, and the time when the boring was performed (before or after the earthquake). To account for these factors while preserving the essential structure of the Seed-Idriss and David-Berrill models the following generalized model is used:

$$\text{logit}(P) = \beta_o + \beta_1 x_1 + \beta_2 x_2 \tag{4.42a}$$

where x_1 would be either ln(CSRN) or ln(Λ) and $x_2 = (N_1)_{60}$ as before. However, the coefficients β_o, β_1 and β_2 are now defined as:

$$\beta_0 = \beta_{00} + \beta_{01}u_1 + \beta_{02}u_2 + \cdots + \beta_{0p}u_p \tag{4.42b}$$

$$\beta_1 = \beta_{10} + \beta_{11}u_1 + \beta_{12}u_2 + \cdots + \beta_{1p}u_p \tag{4.42c}$$

$$\beta_2 = \beta_{20} + \beta_{21}u_1 + \beta_{22}u_2 + \cdots + \beta_{2p}u_p \tag{4.42d}$$

where u_1, u_2 ... u_p are the additional explanatory variables.

Consider a model with only one "primary" explanatory variable x_1 and one "secondary" explanatory variable u_1, where x_1 is continuous but u_1 is

binary and can take on values of 0 or 1. The simplified version of Eqn. 4.42 is in this case:

$$\text{logit}(P) = \beta_o + \beta_1 x_1 \tag{4.43a}$$

or

$$\text{logit}(P) = (\beta_{00} + \beta_{01}u_1) + (\beta_{10} + \beta_{11}u_1)x_1 \tag{4.43b}$$

If $u_1 = 0$, then we obtain the model

$$\text{logit}(P) = \beta_{00} + \beta_{10}x_1 \tag{4.44}$$

On the other hand, if $u_1 = 1$, then

$$\text{logit}(P) = (\beta_{00} + \beta_{01}) + (\beta_{10} + \beta_{11})x_1 \tag{4.45}$$

Thus, the terms $\beta_{01}u_1$ and $\beta_{11}u_1$ may be thought of as changes in β_0 and β_1 resulting from a change of $u_1 = 0$ to $u_1 = 1$.

In Eqn. 4.33, it is conceivable to have either β_{01} or β_{11} equal to zero. If $\beta_{01} = 0$, then the variable u_1 does not affect the intercept coefficient β_0, whereas, in the case where $\beta_{11} = 0$, then u_1 does not affect the slope coefficient β_1. It should be noted that if both β_{01} and β_{11} are non-zero, the logit model will yield the same result as when the two populations (one with $u_1 = 0$ and the other with $u_1 = 1$) are analyzed separately. Thus, incorporating the indicator variable u_1 in both the β_o and β_1 coefficients is equivalent to partitioning of the data according to u_1.

The effects of individual changes of β_0 and β_1 in a single explanatory variable logit model are schematically illustrated in Fig. 4.15. A change in the intercept coefficient β_0 results in a simple translation of the logistic regression curve (Fig. 4.15(a)) whereas a

change in β_1 produces a change of slope of the curve (Fig. 4.15(b)). Fig. 4.15(c) shows the effects of changes in both β_0 and β_1. In two dimensions (i.e. two explanatory variables x_1 and x_2), analogous changes in β_0, β_1, and β_2 may be envisioned as shifting or rotating the equi-probability contour lines or increasing the separation between them.

In terms of computer implementation of Eqn. 4.42, the basic logit regression procedures can be used directly with a simple redefinition of explanatory variables. Using the one variable case as an example, Eqn. 4.43 can be rewritten as

$$\text{logit}(P) = \beta_{00} + \beta_{10}x_1 + \beta_{01}u_1 + \beta_{11}u_1x_1 \qquad (4.46)$$

$$= \beta'_0 + \beta'_1x'_1 + \beta'_2x'_2 + \beta'_3x'_3$$

where the $x'_1 = x_1$, $x'_2 = u_1$, and $x'_3 = u_1x_1$. The terms β'_0, β'_1, β'_2, amd β'_3 would be the corresponding logit coefficients obtained with these redefined variables.

From the preceding discussion, it is seen that the generalized model presented in Eqn. 4.42 is a reasonable way of incorporating the effects of "secondary" variables in the base logit models. The model is simple to implement and to interpret. More complex models involving cross product terms of the secondary variables such as $u_1.u_2$, $u_2.u_3$, etc. are conceivable but were not used because of limitations of the data and the increased complexity of interpretation.

4.7 Significant and Insignificant Secondary Variables

Several secondary variables were considered for possible

incorporation into the Seed-Idriss and Davis-Berrill base models using the format of Eqn. 4.42. The significance and physical implications of those variables are discussed in subsequent sections. The purpose of the present section is to describe the search/selection procedure used to identify those variables that are statistically significant and those that are not.

The term "statistically significant" is used here as a shorthand terminology for the formal statement that the null hypothesis (that a certain variable or model is unimportant) is insignificant at the $\alpha = 0.05$ level and can be rejected. The signficance level $\alpha = 0.05$ was chosen after some initial trials with other α-levels. The results with $\alpha = 0.05$ appeared to be compatible with what was expected in terms of physical behavior of liquefaction, whereas lower α-levels were found to exclude some important variables as being insignificant. Of course, the fact that a variable is statistically significant does not necessarily imply that the same variable is physically meaningful. For example, statistical significance can arise because of biases in the data. Thus, the results of statistical analyses should always be tempered with engineering judgment.

Most of the candidate secondary variables considered in the analysis are binary. Approximately 30 of these were considered in "groups" that were indicators of:

- The source catalog of each case study.
- Geographic location (Japan, California, China, other).
- Whether the epicentral distance was obtained from instrumental data or estimated from intensity isoseismals.
- Whether the distance to energy release is estimated independently

of the epicentral distance.

- How peak site acceleration was obtained.
- The method of performing the SPT.
- How a representative SPT N-value for the case study was obtained.
- Whether a boring was performed before or after the earthquake.

Details of how the above "groups" of variables (and others) have been structured may be found in Appendix A or Chapter 3.

The elimination of statistically insignificant secondary variables was done in two stages. In the first stage, each of the variables belonging to the same "group" were considered and stepwise logistic regression (Sec. 4.3.5) was used to ferret out insignificant variables within the "group". The second stage of the elimination process then used those variables not eliminated during the first stage in another run of the stepwise regression procedure. It was found that this procedure helped to assure physically sensible results. The specific outcomes of this two-stage elimination process are described in Section 4.9.

As mentioned previously in Section 4.3.5, the stepwise logistic regression procedure is a method that results in local rather than a global optimization of the combination of explanatrory variables. In order to verify that the results were at least close to the global optimum, several runs of the stepwise regression algorithm were made using different initial models. Numerous checks of individual logit regression models were also performed. It should be noted that even globally optimum models are not necessarily consistent with physical considerations.

A general overview of the results of search/selection procedure for secondary explanatory variables is shown in Table 4.7 for the Seed-Idriss

and Davis-Berrill base models. Four categories of significance are used in this table to classify the variables. The categories are actually defined as a combination of statistical significance combined with a judgmental consideration of physical interpretation.

It is interesting that the statistical significance of the secondary variables is often different for the Seed-Idriss and Davis-Berrill models. Particularly surprising is the fact that fines content (FC), gravel context (GC) and the median grain size (D_{50}) appear to be important for the Seed-Idriss model, but these factors are statistically insignificant for the Davis-Berrill model. Several conjectures can be made as to why this is so, although it is difficult to reach definitive conclusions, given the current state-of-knowledge of the liquefaction phenomenon and the limitations of the data base.

One conjecture is that the Davis-Berrill model is an inherently "cruder" model of liquefaction behavior than the Seed-Idriss model. If one accepts that the characterization of earthquake intensity (in terms of peak acceleration) for an A&D model has more diagnostic power than an M&R characterization, then this conjecture is a logical conclusion. The argument then follows that grain size effects are too subtle to be distinguished by the Davis-Berrill model (and M&R models in general) though they are significant to the more "refined" Seed-Idriss model. A confirmation of this argument is illustrated in Chapter 7. Another possibility is that A&D models are more subject to prejudices in the assessment of the facts of a case study. Judgement is often required to obtain a value of peak acceleration. If a range of possible accelerations can be inferred, one could be inclined to think that a relatively high acceleration occurred at the site in analyzing a case

study after the fact of liquefaction. On the other hand, at a site where liquefaction did not occur, there would be more of a prejudice towards thinking that the acceleration was relatively low. The evaluation of magnitudes and distances are less subject to interpretation, and thus may partially account for the differences in M&R and A&D models.

Similar arguments could also be made to explain why other secondary variables seem to influence one model and not the other. The time of the boring (before or after the earthquake) is another variable that is perplexing. For this particular variable, firm conclusions cannot be reached due to the relatively small number (39) of borings which are known with certainty to have been performed prior to earthquake occurrence. The possible influence of the sparcity of data on the results for specific secondary variables is also noted in Table 4.7.

The necessity for caution in the interpretation of the results of stepwise logistic regression is further illustrated by the secondary variables that indicate SPT procedure and the method of obtaining site accelerations. It is not clear why the SPT method indicators should be relevant, since differences among the various methods of performing the SPT are accounted for in the data catalog using a normalized N_1-value (i.e. $(N_1)_{60}$) as proposed by Seed et al. (1984). Additional data is also needed to investigate the effects of those variables included in the last category in Table 4.7. These are variables that, on a purely physical basis, <u>should</u> have some effect on probability of liquefaction. For example, non-level ground conditions should be important in many situations. The documentation of additional case studies focussing on the variables indicated in the last significance category of Table 4.7 would

yield useful results.

4.8 Effects of Gradation Variables

The search/selection procedure for statistically significant variables has indicated that three grain-size characteristics are important for liquefaction prediction: fines content (FC), gravel content (GC), and median grain size (D_{50}). In this section, we first consider models that incorporate each of these factors, one at a time, and then models that simultaneously consider two or all three of them. The results presented in this section pertain only to the Seed-Idriss base model. As noted in Section 4.7, the Davis-Berrill model was found to be insensitive to gradation effects.

4.8.1 Fines Content

The effect of fines content (FC) is best modelled by a binary indicator variable denoted as FCI (Fines Content Indicator) where

$$FCI = \begin{cases} 0 & \text{if FC} < 12\% \\ 1 & \text{if FC} > 12\% \end{cases} \tag{4.47}$$

The logit equation using this variable is written as:

$$\text{logit}(P) = \beta_0 + \beta_1 \ln(CSRN) + \beta_2 (N_1)_{60} \tag{4.48}$$

where

$$\beta_0 = \beta_{00} + \beta_{01} \cdot FCI \tag{4.49a}$$

$$\beta_1 = \beta_{10} + \beta_{11} \cdot FCI \tag{4.49b}$$

$$\beta_2 = \beta_{20} + \beta_{21} \cdot FCI \tag{4.49c}$$

Reasons for this functional form are given below.

In the initial stages of model formulation, it seemed desirable to develop a model with the effect of fines content treated as a binary variable, even though fines content is actually a continuous variable. A precedent for doing this was a similar classification of soils as fine silty sands and coarser clean sands as proposed by Seed et al. (1983) based on the median grain size D_{50}. Also, the liquefaction data catalog contained only 159 case studies (of the total 278) where the fines content was actually measured. For the remaining 119 cases, a rough classification into clean and silty sands could usually be made based on the descriptions of the soils presented in the source references. Thus a binary fines content indicator allows the use of a larger data set.

One problem in the definition of a fines content indicator is that of choosing the "critical" value of the fines content that separates "clean" from "silty" sands. This is equivalent to the problem of determining the FC value at which the fines significantly affect the liquefaction behavior of the soil. Seed et al. (1984) suggest a critical FC value as low as 5%, whereas laboratory studies (Kaufman, 1981; Sherif et al, 1983) indicate critical FC values as high as 10% to 30% depending on the interpretation of the results.

Fig. 4.16 shows the modified likelihood ratio index MLRI or $\bar{\rho}^2$ for a series of logit regression fits to the 159 case study data with measured FC values. The models differ in the critical value of fines content chosen to partition the cases into clean and silty sands. Conclusions that can be drawn from Fig. 4.16 are:

1) The use of a binary indicator variable to create a partitioning of the data into clean and silty sand leads to significant

improvement over the base Seed-Idriss model.

2) A critical FC value chosen anywhere from 4% to 12% fines can be used. (See range of high ρ^2 in Fig. 4.16.)

3) The maximum value of MLRI occurs for a critical value of 12% fines content, though the $\bar{\rho}^2$ value (0.4320) does not clearly indicate a statistical advantage over a choice of other FC critical values in the range between 4% and 12%.

Similar conclusions are reached when using the Seed et al. (1984) data catalog with 124 cases, as shown in Fig. 4.17. In this case, the maximum MLRI occurs for a critical value of FC between 7% and 9%. The values of MLRI are exactly the same within this range of critical FC values because the catalog contains no data with fines content between 7% and 9%. However, a marked local maximum in the MLRI occurs for this catalog also at 12% fines content.

Based on the above observations, it was thus decided to choose FC = 12% as the critical value for partitioning the case studies into clean and silty sand subsets. This value has the additional appeal in that it is also the value that divides soil classifications of SM from SP-SM (or SW-SM) within the Unified Soil Classification System.

Using the 159 data subset with measured fines content values, attempts were made to model the effects of fines content as a continuous variable. The use of a binary indicator variable is analogous to that of a step function with respect to fines content. Below a critical value of FC = 12%, the effects of fines are zero, and above FC = 12%, the full effect occurs. Thus the value of 1 for the variable FCI in Eqn. 4.47 or 4.49 may be thought of as a normalized change in liquefaction resistance.

Fig. 4.18 shows various attempts at using smoother function steps to replace FCI. However, as indicated by the values of $\bar{\rho}^2$ associated with the various functions, no significant improvement is obtained with any of the continuous-variable models over the binary-variable formulation.

There are two important conclusions that can be drawn from these results. The first is that modelling the effect of fines content as a binary indicator variable is just as good as modelling it as a continuous variable. The binary variable is also simpler and has the advantage of allowing the use of a larger data set in regresion. Secondly, if there is a "continuous" effect of fines content, the increases in liquefaction resistance due to fines content are already at a maximum once FC exceeds about 12% to 15%. A soil with a certain N_1-value and FC = 30% would have no more additional resistance to liquefaction than a soil with the same N_1-value but with FC = 15%.

The above conclusions appear to contradict the results presented by Tokimatsu and Yoshimi (1983) and Seed et al. (1984). The conclusions are clearly in disagreement with the laboratory results of Kaufman (1981) and Sherif et al. (1983). However, a possible explanation of the apparent contradictions is offered: The SPT is a very crude field test. Unlike refined and controlled laboratory tests, the SPT is fairly insensitive to the fines content of the soil. The SPT may be able to distinguish roughly between clean sands and silty sands, but it cannot distinguish between a sand, for instance, that has 15% fines and a sand with 30% fines. Thus, variations in the fines content may actually influence the liquefaction resistance of a soil, but this influence is not reflected by and cannot be predicted from the SPT N-value.

The logit equation fitted to 278 case studies (182 cases of clean sand, and 96 cases of silty sand) is reported in Table 4.8. The implications of introducing a fines content indicator (FCI) in the Seed-Idriss base model are shown in Figures 4.19 and 4.20. Note that the form of Eqn. 4.49 makes the use of FCI equivalent to fitting the logit equation separately to the two data sets (clean sands and silty sands) as was done in Sec. 4.3 (Table 4.3). The equi-probability contour lines in Fig. 4.19(b) are more dispersed than those in Fig. 4.19(a), reflecting greater uncertainty in the predictive value of the Seed-Idriss model for silty sands. This is probably due to the larger diversity of soils included in the silty sand class.

Fig. 4.20 shows a comparison of the 0.1, 0.5, and 0.9 equi-probability contour lines for the two types of sands. The P = 0.5 contour line of the silty sands is higher than the corresponding line for clean sands, indicating in general that the effect of higher fines content (FC > 12%) is to increase the liquefaction resistance. Thus for silty sands and clean sands with the same N_1-value, the use of the P = 0.5 contour line in a design situation would indicate a <u>higher</u> CSRN value required to liquefy the silty sand. However, if we wished to design for a relatively low (more conservative) conditional probability of liquefaction, say P = 0.1, the results indicate a <u>lower</u> value of CSRN required to liquefy the silty sand. Thus, it would be incorrect to make the general statement, that given the same N_1-value, a silty sand has a lower probability of liquefaction than a clean sand. The reason why such a statement is not always correct is that greater uncertainty exists in predicting the liquefaction behavior of silty sands compared to that of

clean sands. It should also be noted that the assessment of low values of P represent mainly extrapolation, and these should be considered less reliable.

4.8.2 Gravel Content

The effects of gravel content (GC) on the Seed-Idriss and Davis-Berrill base models were also found to be best modelled by a binary indicator variable denoted GCI and defined as:

$$GCI = \begin{cases} 0 \text{ if } GC \leq 10\% \\ 1 \text{ if } GC > 10\% \end{cases} \tag{4.50}$$

The form of the logit equation proposed to incorporate this variable in the Seed-Idriss base model is:

$$\text{logit}(P) = \beta_0 + \beta_1 \ln(CSRN) + (\beta_{20} + \beta_{21} \cdot GCI)(N_1)_{60} \tag{4.51}$$

In contrast to the fines content, the gravel content only affects the coefficient of $(N_1)_{60}$.

The decision to use Eqn. 4.50 and Eqn. 4.51 was arrived at by following a procedure similar to that described in Sec. 4.3.1 for fines content. The results of a series of logit regression analysis using various critical gravel content values for partitioning the data into "gravelly" and "non-gravelly" soils are shown in Fig. 4.21. The results indicate a significant increase in $\bar{\rho}^2$ for a model that accounts for gravel content compared to a model that does not. The maximum value of $\bar{\rho}^2$ occurs for a critical value of GC = 10%, and thus this value was incorporated into the model of Eqn. 4.50. Various attempts were also made to model the effects of gravel content as a continuous variable, but

the goodness-of-fit statistics for these models are no better than when gravel content is included through a binary indicator variable.

The lines of equal probability of liquefaction for the Seed-Idriss model extended to account for the gravel content are shown in Fig. 4.22 and Fig. 4.23. The corresponding fitted logit regression equation is reported in Table 4.8. Basically, the effect of gravel content on the equi-probability contour lines is to rotate them with respect to the contours for non-gravelly sand. The P = 0.1, 0.5, and 0.9 contour lines for gravelly sands and non-gravelly sands are compared in Fig. 4.23. This figure indicates that, particularly for high values of $(N_1)_{60}$, gravelly soils are generally more susceptible to liquefaction than non-gravelly soils.

That gravelly soils may be more susceptible to liquefaction than non-gravelly soils can be explained by the fact that gravelly soils can have higher SPT resistance without a corresponding increase in soil density or liquefaction resistance. In gravelly soils, it is not uncommon for one or two large pieces of gravel to become lodged in the tip of the SPT sampler and hence bias the SPT N-value upwards. As the size of the soil particles becomes comparable to the diameter of the sampler, effects such as these are more likey to occur. Thus, for gravelly soils, the SPT N-values are not representative of the true liquefaction or penetration resistance.

An unfortunate feature of the present data set is that there are only 13 case studies (out of 40 cases where GC $\neq$ 0) involving soils with gravel content GC > 10%. Thus, even though incorporating the effects of gravel content produces better goodness-of-fit statistics, the results

must be viewed as preliminary. It is necessary to collect and process additional data with different gravel content before reaching definitive conclusions.

4.8.3 Median Grain Size

In contrast to the binary indicator variables used to incorporate the secondary effects of fines and gravel content, the median grain size was modelled as a continuous variable. The logistic equation for this purpose is of the form:

$$\begin{aligned} \text{logit}(P) = {} & \beta_{00} + \beta_{01} \ln(D_{50}) \\ & + [\beta_{10} + \beta_{11} \ln(D_{50})] \ln(\text{CSRN}) \\ & + \beta_2 (N_1)_{60} \end{aligned} \qquad (4.52)$$

where D_{50} is the median grain size in millimeters. This form of the equation was arrived at after various runs using stepwise logistic regression procedures and other attempts to find models with improved goodness-of-fit statistics. The β coefficients fitted to the 165 cases with available D_{50} data were shown in Table 4.8.

The effects of varying D_{50} size in the Seed-Idriss model is illustrated in Fig. 4.24, where three sets of contour lines for the probability of liquefaction are shown (for D_{50} = 0.1, 0.3, and 1.0). It is interesting that the effect of D_{50} embodies effects similar to that of both fines and gravel content. As the soils become coarser and cleaner (as D_{50} increases) the dispersion of the equi-probability contour lines decreases. This is similar to the effect of having FC < 12% (see Fig. 4.19). Also, as D_{50} increases, there is a general decrease in liquefaction resistance, as was noticed for the case when GC > 10% (see

Fig. 4.22).

Fig. 4.25 shows a comparison of the P = 0.1, 0.5 and 0.9 equiprobability values for the same three D_{50} sizes used in Fig. 4.24. Based on this comparison. a fairly general conclusion is that, at least for values of $(N_1)_{60}$ larger than 20, there is an increase in liquefaction susceptibility as D_{50} increases. This is in agreement with results presented by Iwasaki et al. (1978) and Tatsuoka et al. (1980), and with provisions in the Japanese Code of Bridge Design (see Fig. 1.2).

4.8.4 Combined Effects of Gradation Variables

The previous sections have presented the results of logitistic regression considering separately the effects of fines content, gravel content and median grain size. In this section, models are analyzed considering various combinations of the variable FCI, GCI and D_{50}. Two models are discussed in detail:

- A model that includes FCI, GCI and D_{50} and is fitted to 165 cases for which D_{50} is available.
- A model that includes FCI and GCI. This model has the advantage that all 278 cases can be used in the regression fitting.

The parametric forms and the fitted logit coefficient for these two models are given in Table 4.9.

The results obtained using the model considering FCI and GCI are illustrated in Fig. 4.26. Basically, the model divides the soils into three groups: silty sand, clean sand, and gravelly sand. These three groups should be considered to be mutually exclusive. Specifically, it would be inappropriate to set GCI = 1 and FCI = 1 in the logit equations

in Table 4.9, since there are no cases in the data set where a silty and gravelly sand was encountered. Because of this reason, the model exhibits the same features of gradation effects as when the indicator variables are included separately in the Seed-Idriss models.

Similar results are illustrated in Fig. 4.27 for a model combining the three variables FCI, GCI, and D_{50}. The contour lines in Fig. 4.27 are for D_{50} = 0.30 mm. The effect of varying D_{50} is shown in Fig. 4.28 where the P = 0.5 contours for clean sand (non-silty and non-gravelly) are plotted for D_{50} = 0.1, 0.3, and 1.0. This figure indicates that, when the fines content and gravel content are accounted for, the effect of increasing D_{50} is to increase liquefaction resistance. Note that this is the direct opposite implication of the model considering D_{50} effects only (see Section 4.8.3). However, this conclusion is consistent with the results of Lee and Fitton (1969).

That models with combined gradation variables are statistically superior to the models that consider only one such variable or no gradation variables at all, is demonstrated in Table 4.10. In order to avoid the effect of differences arising from using different data sets, the statistics presented in this table are all based on the same data. Note in particular that $\bar{\rho}^2$ is quite high for the combined FCI, GCI and D_{50} model.

It is generally recognized that the liquefaction resistance of a soil is dependent on its grain size characteristics. Soils normally considered most susceptible to liquefaction are usually cohesionless fine to medium sands with very low values of fines content (say FC < 12%). In early studies involving laboratory tests (e.g. Lee and Fitton, 1969), the

median grain size D_{50} was often used as an index of the entire grain size curve of the soils used in testing. Seed, Idriss, and Arango (1983) used the D_{50} size to separate soils into two groups ($D_{50} < 0.15$ mm and $D_{50} > 0.25$ mm) for purposes of liquefaction assessment, with finer soils indicated as having more resistance to liquefaction for a given SPT N_1-value. Recently, the trend in laboratory studies has been toward a characterization of liquefaction-susceptible soils based on the fines content rather than the D_{50} size (Kaufman, 1981, Sherif et al, 1983). A similar shift in thinking is also evident in evaluations involving SPT data (Tokimatsu and Yoshimi, 1983; Seed et al., 1984).

Inasmuch as D_{50} and the fines content are correlated, use of either FC or D_{50} can be justified. From Table 4.10, however, it can be seen that in a catalog where both variables are available, the model with fines content (FCI) is superior to the D_{50} model. There is also some correlation between gravel content GC and D_{50}, and between GC and FC (albeit an exclusivity type of relationship in the latter case). To use the variables FCI and GCI together in a model is logical because these two factors contain different pieces of "information" about the soil. To also incorporate D_{50} size as a third variable is, however, somewhat redundant. Indeed, the range of the $P = 0.5$ contour lines shown in Fig. 4.28 is what one might expect from using different data sets (see Figures 4.10, 4.12 and 4.14). In practical terms, this means that the additional precision obtained from considering D_{50} in such a model may be completely offset by imprecisions due to other factors. Thus a good model appears to be one in which FCI and GCI are included. However, this model suffers from the lack of sufficient data on a range of gravelly soils in the

liquefaction catalog, as was mentioned earlier in Section 4.8.2.

The important conclusion that can be derived from the results of this section is that a single grain size parameter (e.g. FC, D_{50}, etc.) may not be sufficient to completely characterize liquefaction resistance. For example, a model which includes D_{15}, D_{50}, and D_{85} may be more appropriate than one which just includes D_{50}. Toward this end, it would be helpful if the entire grain size distribution of the soil were documented in reporting future liquefaction or non-liquefaction cases.

4.9 Effects of Other Variables

As mentioned previously in Section 4.7, several binary indicator variables other than those associated with soil gradation effects were considered for candidate secondary variables. In this section, we report results of the stepwise regression analysis for the Seed-Idriss model (based on all 278 cases of the liquefaction catalog) and for the Modified (hypocentral) Davis-Berrill Model (based on 228 cases).

The results presented in this section are much more tentative than those presented in the previous sections. This is due to two reasons. The first is that many of the subsets of data created by the use of binary variables are relatively small. For example, only 16 case studies were documented where a site response analysis was performed to estimate acceleration. Based on logit analysis, it appears that these cases present an inhomogeneity within the liquefaction catalog, but additional data might alter such a conclusion obtained at this time from a limited sample size. The second reason is due to the difficulty of explaining the physical implications of some results. The present analyses are

therefore presented mainly for the purpose of exploring the data and pointing out anomalies rather than reaching definite conclusions.

4.9.1 Seed-Idriss Model

A logit equation with seven explanatory variables obtained by using all 278 cases is given in Table 4.11. Contours of equal probability of liquefaction for this model are shown in Fig. 4.29 and Fig. 4.30 with various indicator variables activated (i.e. set equal to 1). Fig. 4.29(a) shows the "base" model for clean sand to which the remaining parts of the figure should be compared. Similarly, Fig. 4.30(a) shows the "base" model for silty sand. Note in Fig. 4.30(c) the clearly incorrect shapes of the contour lines. This figure is presented to show that certain subsets of the data contain anomalies, and for them, the model presented is not acceptable.

Figures 4.31 and 4.32 show P = 0.5 contours respectively from Figures 4.29 and 4.30 plotted for comparison. The effects of the two gradation indicator variables are the same as those observed in Section 4.8. In addition the figures indicate that:

- Documented sites from regions other than Japan, California, or China are generally much less susceptible to liquefaction, at least for the clean sand subset.
- Cases where the acceleration was estimated from site response analysis are slightly less susceptible to liquefaction. One explanation might be that response analysis tends to overestimate the acceleration at the site.
- Cases which involved accelerations estimated from intensity data (and/or some attentuation relationship other than those proposed

by Joyner and Boore (1981) or by Kawashima et al. (1984)) appear to be slightly more susceptible to liquefaction, i.e. perhaps the accelerations are slightly underestimated. [See Chapter 3 for details of various methods.]

- Cases for which the SPT was performed using the rope and pulley technique are more susceptible to liquefaction, i.e. the SPT rope-and-pulley method over-estimates the penetration resistance compared to the "free-fall" or trip-hammer method.

Several physical reasons could be postulated for the trends described above. However, those trends may just as well be due to the inadequacy of the data catalog.

4.9.2 Modified Davis-Berill Model

Stepwise logistic regression procedures identified fewer significant secondary variables for the hypocentral Davis-Berrill model than for the Seed-Idriss model. The two variables found to be significant are those associated with the method of analyzing a boring log to obtain a "representative" N-value (see Chapter 3) and with the time (before or after the earthquake) when the boring was made.

The logit equation fitted to the 228 data subsets with reported focal depths is given in Table 4.12. Fig. 4.33 shows the implications of this logit model. As indicated in the figure, if the method of obtaining the "representative" SPT N-value is unknown, there is a slight rotation and shift of the contour lines of liquefaction probability. This indicates that in these cases it is likely that the SPT value was overestimated. This conclusion is consistent with what was known to have been done in most cases documented in the catalog, which was to choose the minimum

N_1-value along the boring profile as the "representative" value.

Another implication of the fitted logit model is that the N-values obtained in a boring tend to increase after the occurrence of an earthquake. This is indicated by Fig. 4.33(c) where the equi-probability lines are rotated counterclockwise and with respect to those shown in Fig. 4.33(a). Note that this shift is consistent with generally accepted notions of physical behavior. However, the actual physical behavior has not been clearly documented, though attempts have been made to do so (e.g. Koizumi, 1966). Youd (1984) has proposed that previously liquefied sites are also more prone to liquefy in the future. However, this statement and the notion that N-values tend to increase after earthquake shaking are not incompatible. Sites may re-liquefy with successive earthquakes, but the effects of liquefaction become successively less severe. This has been documented by the data from centrifuge experiments performed by Heidari (1981).

Compared to the non-gradation secondary effects in the Seed-Idriss model, the secondary effects described here are more amenable to physical interpretation. However because of the limited data base (e.g. only 39 case study borings were definitely performed prior to the earthquake) caution must again be exercised in directly utilizing these logit regression results.

4.10 Chapter Summary

The methodology of binary logistic regression has been presented and used to analyze the liquefaction data catalog compiled for this study. Two models -- one based on the A&D type Seed-Idriss formulation, and the other

based on the M&R type Davis-Berrill formulation -- have been established and were shown to fit the data fairly well. Statistically signficant secondary factors such as soil gradation measures, SPT method, etc. have been considered for incorporation with both models. However, the secondary variables that are statistically significant for the Seed-Idriss model differ from those that are significant for the Davis-Berrill model.

The Seed-Idriss model was chosen for analysis because of its prevalent use by the geotechnical engineering profession. From among several candidate M&R models, the Davis-Berrill model was chosen based on its goodness-of-fit to the data and on physical considerations. As originally formulated, the Davis-Berrill model is in terms of the epicentral distance. However, a modified Davis-Berrill model which uses the hypocentral distance gives better results was also used in this study. Attempts at formulating models based on the distance to energy release (DER) were unsuccessful. At present, the data cannot be used to justify an M&R model based on DER.

The proposed logit model incorporating the Seed-Idriss method of analysis is shown to give concave upward shapes for the contour lines of equal probability of liquefaction. However, the contour lines (or any discriminant line separating liquefaction and non-liquefaction regions) do not necessarily pass through the origin of the CSRN and $(N_1)_{60}$ axes. The results imply that even at very low blow counts ($N \approx 0$), a threshold value of CSRN is required to cause liquefaction. This is hypothesized to be the combined result of the critical strain concept (Dobry et al, 1982) and the procedures involved in performing the SPT.

The effect of soil gradation on the Seed-Idriss base model has been

investigated. Three factors that are statistically significant are the fines content (FC), the gravel content (GC), and the median grain size (D_{50}). Considering each of these factors in separate logistic models have yielded the following trends:

- Soils with FC > 12% can be considered as being usually less susceptible to liquefaction, but there is a great deal of uncertainty in this statement, particularly if a conservative design is desired.
- Soils with GC > 10% are more susceptible to liquefaction than soils with GC < 10% for the same SPT resistance.
- Increasing D_{50} causes an increase in liquefaction susceptibility for soils with the same SPT resistance.

In a combined model incorporating all these variables, the trends are the same as above, except for the effect of D_{50}.

Soil gradation characteristics have statistically insignificant effects on the Davis-Berrill model. One possible reason is that the Davis-Berrill model and other M&R models do not have as much diagnostic power as A&D models. This, of course assumes that a "local" A&D formulation is the proper characterization, for which "global" M&R models are an approximation. However, this is not an argument against employing M&R models in risk analysis, where these models may actually have an advantage in that the uncertainty of earthquake attenuation is already incorporated, whereas for A&D models this uncertainty is not.

Other qualitative conclusions have been reached by considering different indicator variables. However, these conclusions are tentative largely because of limitations of the available data. One particular

variable that deserves to be mentioned is the indicator of whether the case study boring was performed before or after the earthquake. This factor is signficant in the Davis-Berrill model and indicates that an increase in SPT resistance is expected for borings performed after earthquakes.

The main results presented in this chapter are insensitive to the data base used for analysis. This statement is based on logit regression analyses performed on the Seed et al. (1984) data and on various subsets of the catalog compiled for this study.

Table 4.1

Examples of the Use of Logit Analyses in other Disciplines

PROBLEM	EXPLANATORY VARIABLE(S)	RESPONSE VARIABLE
Effectiveness of pesticides	Dosage of pesticide	Death/Survival
Effectiveness of drugs	Dosage of drug	Cure/No cure
Determinants of coronary heart disease	Serum cholesterol levels, blood pressure, age, and sex of individuals	Heart disease/ No heart disease
Determinants of breast cancer in in women	Number of children, history of hysterectomy, previous benign tumors, duration of symptoms	Cancer/No cancer
Voting habits of individuals	Party affiliation, family income, age and sex of individuals	Candidate X/ Candidate Y
College going behavior	SAT scores, family income, sex of individual	Attend college/ Not attend college
Choice of transportation mode in commuting	Distance to job location, time and cost of travel for each mode, income of individual	Auto/Bus/Train

Table 4.2

Logit Regression Results of Models Fitted Using
Various Parametric Forms - All 278 Data Cases

MODEL*	FITTED LOGIT EQUATION	PCP	MLRI
x_1 = CSRN x_2 = $(N_1)_{60}$	logit(P) = -1.8490 + 25.741 CSRN - 0.25666 $(N_1)_{60}$	82.4	0.4446
x_1 = CSRN x_2 = ln[$(N_1)_{60}$]	logit(P) = 0.44885 + 21.615 CSRN - 1.9950 ln[$(N_1)_{60}$]	78.4	0.3912
x_1 = ln(CSRN) x_2 = $(N_1)_{60}$	logit(P) = 10.167 + 4.1933 ln(CSRN) - 0.24375 $(N_1)_{60}$	83.1	0.4844
x_1 = ln(CSRN) x_2 = ln[$(N_1)_{60}$]	logit(P) = 12.218 + 4.0020 ln(CSRN) - 2.2856 ln [$(N_1)_{60}$]	82.4	0.4492

*Note: Models fitted to the general logit equation:

$$\text{logit}(P) = \ln[P/(1-P)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

or equivalently

$$P = 1/[1 + \exp\{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2)\}]$$

Table 4.3

Logit Regression Result of Models Fitted Using Various Parametric Forms - Data Cases Partitioned into Clean and Silty Sand Subsets

MODEL*	FITTED LOGIT EQUATION	PCP	MLRI
"Clean" Sand Subset of 182 Cases with Fines Content FC < 12%			
$x_1 = \ln(CSRN)$ $x_2 = (N_1)_{60}$	$\text{logit}(P) = 16.447 + 6.4603 \ln(CSRN) - 0.39760 (N_1)_{60}$	85.7	0.6337
$x_1 = \ln(CSRN)$ $x_2 = \ln[(N_1)_{60}]$	$\text{logit}(P) = 24.886 + 7.2036 \ln(CSRN) - 5.0281 \ln(N_1)_{60}$	86.8	0.6275
"Silty" Sand Subset of 96 Cases with Fines Content FC > 12%			
$x_1 = \ln(CSRN)$ $x_2 = (N_1)_{60}$	$\text{logit}(P) = 6.4831 + 2.6854 \ln(CSRN) - 0.18190 (N_1)_{60}$	70.8	0.2652
$x_1 = \ln(CSRN)$ $x_2 = \ln[(N_1)_{60}]$	$\text{logit}(P) = 7.5524 + 2.5994 \ln(CSRN) - 1.4413 \ln[(N_1)_{60}]$	67.7	0.2552

*Note: Refer to text or Note on Table 4.2 for model notation.

Table 4.4

Comparison of Proposed M&R Load Functions in Liquefaction Analysis

MODEL	LOAD FUNCTION
Yegian and Whitman (1978)	$S_c = \frac{10^{0.217M}}{R_{HY} + 16} \frac{\sigma_v}{\bar{\sigma}_v}$
Yegian and Vitelli (1981a, 1981b)	$S_c = \frac{10^{0.0868M}}{R_{HY} + 16} \frac{\sigma_v}{\bar{\sigma}_v}$
Davis-Berrill (Epicentral Model)	$\Lambda_{EP} = \frac{10^{1.5M}}{(R_{EP})^2 (\bar{\sigma}_v)^{1.5}}$
Modified Davis-Berrill (Hypocentral Model)	$\Lambda_{HY} = \frac{10^{1.5M}}{(R_{HY})^2 \bar{\sigma}_v^{1.5}}$

Notation and Units:

M = Earthquake (Richter) magnitude

R_{EP} = Epicentral distance (km)

R_{HY} = Hypocentral distsance (km)

σ_v = Total overburden stress (kg/cm^2)

$\bar{\sigma}_v$ = Effective overburden stress (kg/cm^2)

Table 4.5

Comparison of Various M&R Logit Regression Models

MODEL	FITTED LOGIT EQUATION	PCP	MLRI
Generalized M&R Epicentral Model	$\text{logit}(P) = -13.702 + 3.0976\,M - 1.6217\,\ln(R_{EP}) - 1.3201\,\ln(\bar{\sigma}_v) - 0.23130\,(N_1)_{60}$	81.6	0.4364
Generalized M&R Hypocentral Model	$\text{logit}(P) = -14.736 + 3.8524\,M - 2.6083\,\ln(R_{HY}) - 1.3761\,\ln(\bar{\sigma}_v) - 0.23200\,(N_1)_{60}$	82.5	0.4677
Yegian and Whitman (1978)	$\text{logit}(P) = -1.3639 + 2.1951\,\ln(S_c) - 0.14458\,(N_1)_{60}$	72.8	0.2522
Yegian and Vitelli (1981a, 1981b)	$\text{logit}(P) = 1.0352 + 3.3281\,\ln(S_c) - 0.14554\,(N_1)_{60}$	72.4	0.2326
Davis-Berrill (Epicentral Model)	$\text{logit}(P) = -13.239 + 0.90652\,\ln(\Lambda_{EP}) - 0.22905\,(N_1)_{60}$	81.6	0.4403
Modified Davis-Berrill (Hypocentral Model)	$\text{logit}(P) = -15.143 + 1.0837\,\ln(\Lambda_{HY}) - 0.22656\,(N_1)_{60}$	82.5	0.4664

Note: Models are fitted using 228 case studies that have available data on earthquake focal depth.

Table 4.6

Compairson of Various Forms of the
Generalized Load Function Ψ

MODEL	FITTED OR INFERRED LOAD FUNCTION
Generalized M&R Epicentral Model	$\Psi = \frac{10^{1.35M}}{(R_{EP})^{1.62}(\bar{\sigma}_v)^{1.32}}$
Generalized M&R Hypocentral Model	$\Psi = \frac{10^{1.67M}}{(R_{HY})^{2.61}(\bar{\sigma}_v)^{1.38}}$
Yegian-Whitman	$\Psi = \frac{10^{0.476M}}{(R_{HY}+16)^{2.195}}\left[\frac{\sigma_v}{\bar{\sigma}_v}\right]^{2.195} = S_c^{2.195}$
Yegian-Vitelli	$\Psi = \frac{10^{0.289M}}{(R_{HY}+16)^{1.331}}\left[\frac{\sigma_v}{\bar{\sigma}_v}\right]^{3.328} = S_c^{3.328}$
Davis-Berril (Epicentral Model)	$\Psi = \frac{10^{1.36M}}{(R_{EP})^{1.81}(\bar{\sigma}_v)^{1.35}} = (A_{EP})^{0.906}$
Modified Davis-Berrill (Hypocentral Model)	$\Psi = \frac{10^{1.62M}}{(R_{HY})^{2.16}(\bar{\sigma}_v)^{1.62}} = (A_{HY})^{1.08}$

Table 4.7

Statistical Significance of Secondary Factors or Variables Affecting Liquefaction Occurrence

SIGNIFICANCE CATEGORY	FACTORS OR VARIABLES	
	SEED-IDRISS MODEL	DAVIS-BERRILL MODEL
Statistically significant with definitely meaningful physical interpretation	• Fines content • Gravel content • Median grain size	• Use of hypocentral distance in the model • Boring performed prior to earthquake‡
Statistically significant with possibly meaningful physical interpretation	• Data other than from Japan/ California/ China‡ • Method of obtaining site acceleration • Method of SPT	• Method of obtaining representative N-value • Availability of data for distance to energy release
Possibly statistically significant as indicated by preliminary variable selection procedures	• Method of obtaining representative N-value • Association with cases specifically evaluated for this study	• Gravel content‡ • Japan/California/China /Other regional distinctions • Method of obtaining site acceleration
Statistically insignificant, but may be obscured by lack of data or lack of knowledge	• Boring performed prior/after earthquake‡ • Non-level ground‡ • Artesian conditions • Uniformity coefficient • Clay content	• Effects of gradation • Non-level ground‡ • Artesian conditions‡ • Use of distance to energy release as a distance measure

Note: (‡) indicates a scarcity of data related to these variables that may influence the results shown in this table.

Table 4.8

Results of Logit Analysis Using the Seed-Idriss
Model Accounting for Various Individual Gradation Effects

SECONDARY VARIABLES/ DATA BASE	FITTED LOGIT EQUATION	PCP	MLRI
Fines Content Indicator $FCI = \begin{cases} 0 & FC < 12\% \\ 1 & FC > 12\% \end{cases}$ Using all 272 case studies	$\text{logit}(P) = 16.447 - 9.9640 \cdot FCI + (6.4603 - 3.7750 \cdot FCI)\ \ln(CSRN) - (0.39760 - 0.21570 \cdot FCI)(N_1)_{60}$	80.6	0.5065
Gravel Content Indicator $GCI = \begin{cases} 0 & GC < 10\% \\ 1 & GC > 10\% \end{cases}$ Using all 278 case study data	$\text{logit}(P) = 10.324 + 4.1961\ \ln(CSRN) - (0.26619 - 0.10186 \cdot GCI)(N_1)_{60}$	84.2	0.4853
Median Grain Size D_{50} using 165 case subset with measured D_{50}(mm)	$\text{logit}(P) = 15.607 + 2.7339\ \ln(D_{50}) + [7.0842 + 1.4219\ \ln(D_{50})]\ \ln(CSRN) - 0.25345\ (N_1)_{60}$	81.2	0.4483

Table 4.9

Results of Logit Analysis Using the Seed-Idriss
Model Accounting for Combined Gradation Effects

SECONDARY VARIABLES/ DATA BASE	FITTED LOGIT EQUATION	PCP	MLRI
Fines content indicator FCI Gravel Content Indicator GCI All 278 case study data	$\text{logit}(P) = 16.932 - 10.291 \cdot FCI + (6.5268 - 3.8042 \cdot FCI)\ln(CSRN) + [0.44038 - 0.2420 \cdot FCI - 0.11335 \cdot GCI](N_1)_{60}$	80.9	0.5157
Fines Content Indicator FCI Gravel Content Indicator GCI Median Grain Size D_{50} (mm) 165 cases w/D_{50} Data	$\text{logit}(P) = 17.226 - 11.578 \cdot FCI - 1.1138 \cdot \ln(D_{50}) + (7.2491 - 3.6030 \cdot FCI)\ln(CSRN) - [0.46026 - 0.27794 \cdot FCI - 0.26070 \cdot GCI](N_1)_{60}$	82.4	0.4966

Table 4.10

Comparison of Goodness-of-Fiot of Models with Various Combinations of Gradation Effect Variables

SECONDARY VARIABLES	PCP	MLRI
None (Base Model)	80.0	0.4395
FCI	80.0	0.4588
GCI	81.2	0.4562
D_{50}	81.2	0.4483
FCI, D_{50}	78.8	0.4693
FCI, GCI	80.6	0.4760
FCI, GCI, D_{50}	82.4	0.4966

Note: Statistics are for 165 case subset with D_{50} data.

Table 4.11

Logit Analysis Results for Seed-Idriss Model
Incorporating Signficant Indicator Variables
(fitted to all 278 data)

Model: $\text{logit}(P) = \beta_o + \beta_1 \ln \text{CSRN} + \beta_2 (N_1)_{60}$

Coefficients:

$\beta_o = 25.036 - 15.553 \cdot \text{FCI} - 12.888 \cdot \text{REGO} + 1.7128 \cdot \text{AC4} + 1.9244 \cdot \text{SPT1}$

$\beta_1 = 10.048 - 6.2054 \cdot \text{FCI} - 5.3209 \cdot \text{REGO} + 2.2498 \cdot \text{AC3}$

$\beta_2 = -0.62161 + 0.35453 \cdot \text{FCI} + 0.18773 \cdot \text{GCI}$

Indicator Variables:

$$\text{FCI} = \begin{cases} 1 & \text{if FC} > 10\% \\ 0 & \text{if FC} < 10\% \end{cases} \qquad \text{GCI} = \begin{cases} 1 & \text{if GC} > 10\% \\ 0 & \text{if GC} < 10\% \end{cases}$$

$$\text{REGO} = \begin{cases} 1 & \text{if case from Japan, California, or China} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{AC3} = \begin{cases} 1 & \text{if acceleration based on response analysis} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{AC4} = \begin{cases} 1 & \text{if acceleration based on intensity data (see text for more details)} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{SPT1} = \begin{cases} 1 & \text{if Rope \& Drum Method used for SPT} \\ 0 & \text{otherwise} \end{cases}$$

Goodness-of-Fit Statistics: PCP = 88.1 MLRI = 0.6125

Table 4.12

Logit Analysis Results for Modified Davis-Berrill Model Incorporating Significant Indicator Variables (Fitted to 228 Data)

Model: $\text{logit}(P) = \beta_0 + \beta_1 \ln \Lambda_{HY} + \beta_2 (N_1)_{60}$

Coefficients:

$\beta_0 = -16.925$

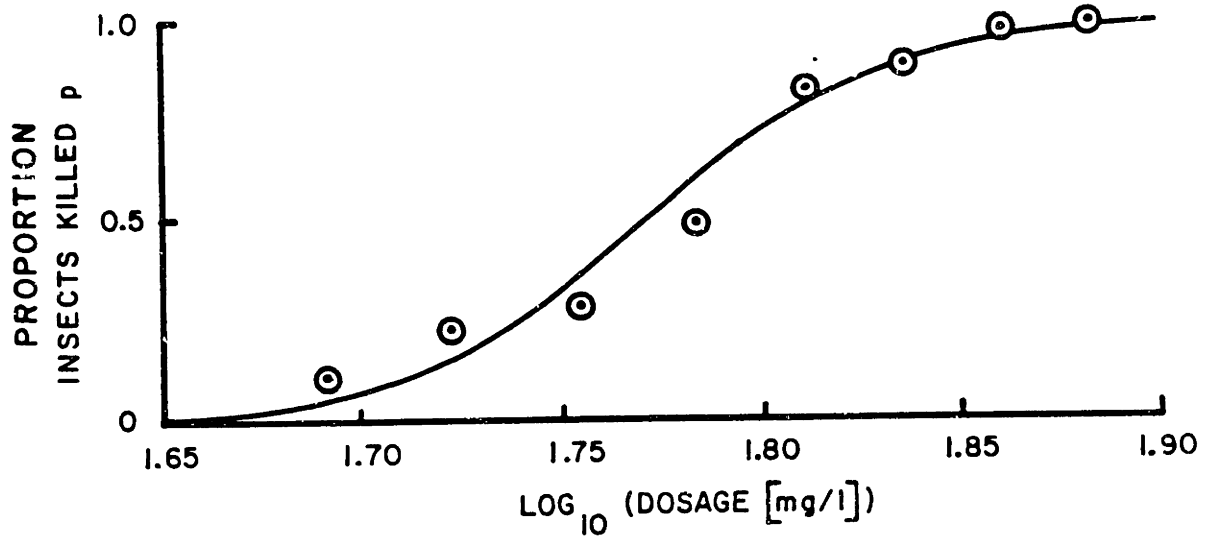
$\beta_1 = 1.1496 + 0.093477 \cdot TB1$

Indicator Variables:

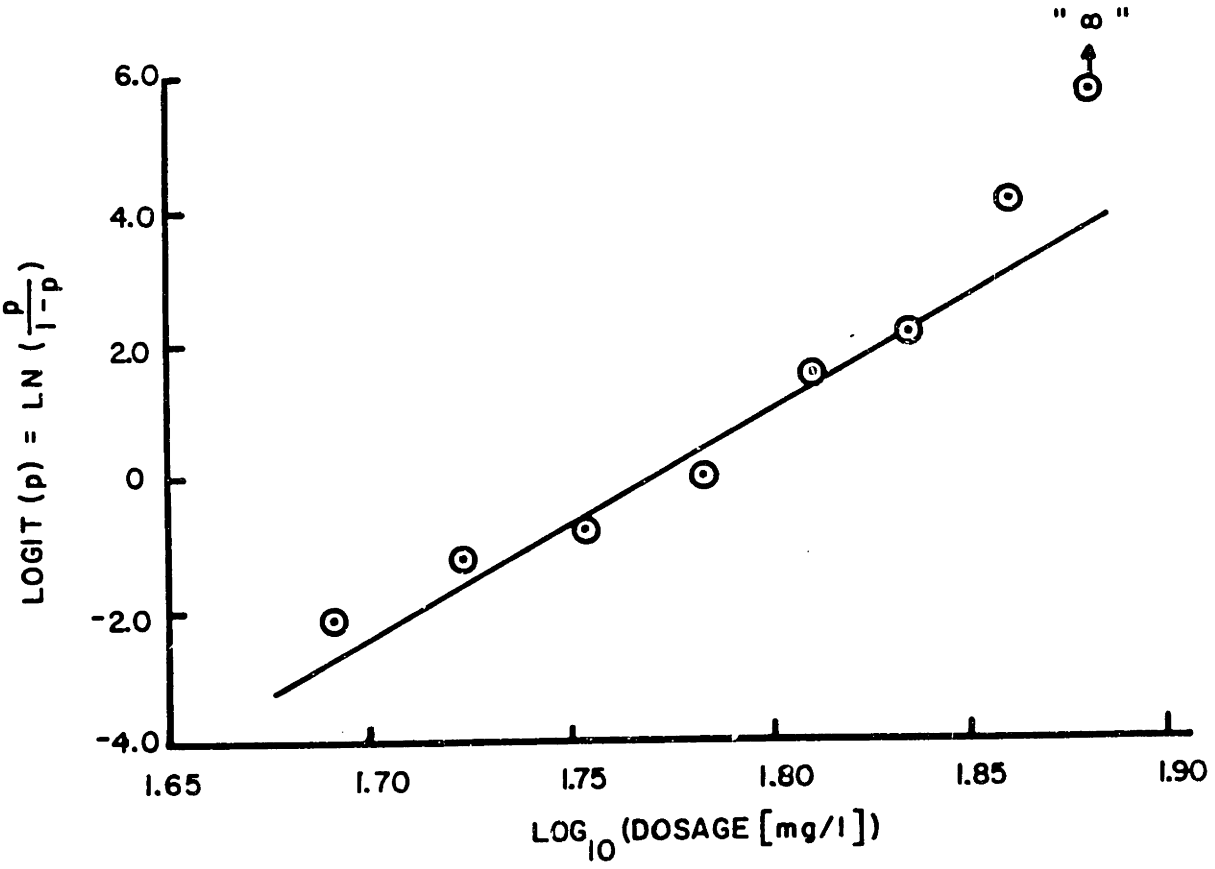
$$NAVO = \begin{cases} 1 & \text{if method of selecting representative N-value is unknown} \\ 0 & \text{otherwise} \end{cases}$$

$$TB1 = \begin{cases} 1 & \text{if case study boring is positively known to be performed prior to earthquake} \\ 0 & \text{otherwise} \end{cases}$$

Goodness-of-Fit Statistics: PCP = 85.1 MLRI = 0.5207



(a)



(b)

Fig. 4.1 Illustration of the logit transform applied to a dose-response experiment. The data are from Bliss (1935) for beetles exposed to gaseous carbon disulfide. The curve and the lines shown are maximum likelihood fits reported by Dobson (1983).

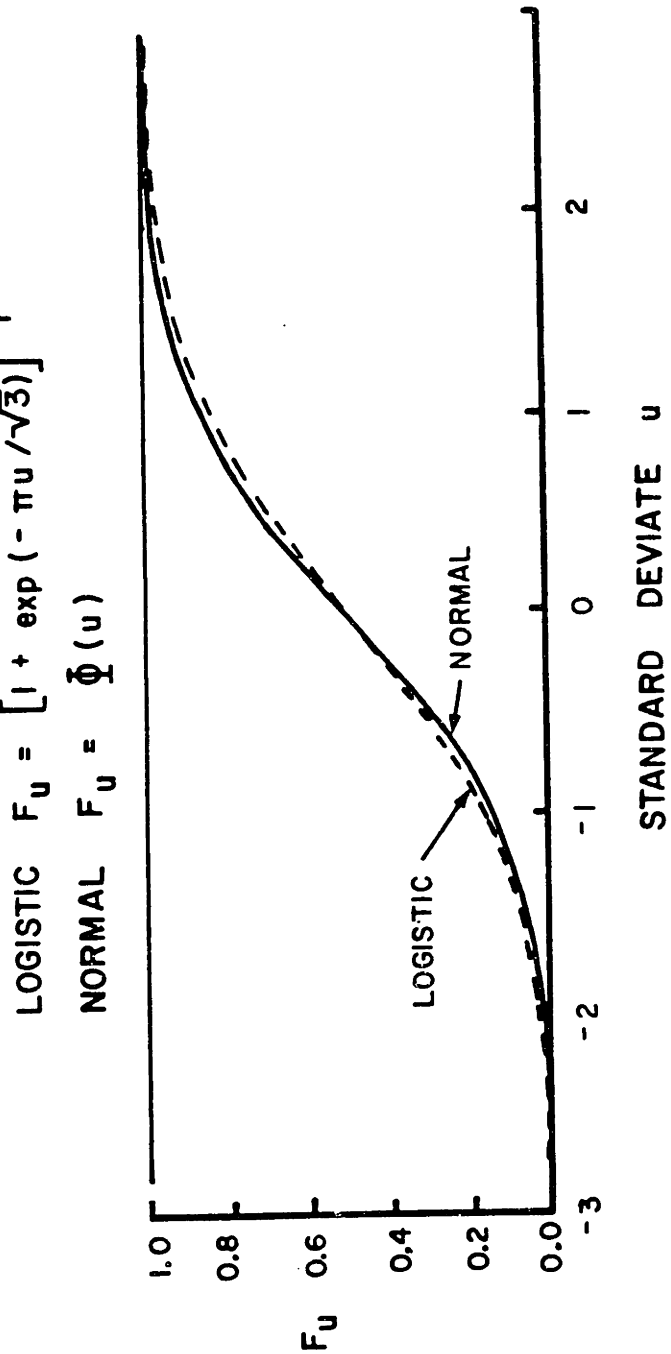


Fig. 4.2 Comparison of logistic and normal cumulative distribution functions.

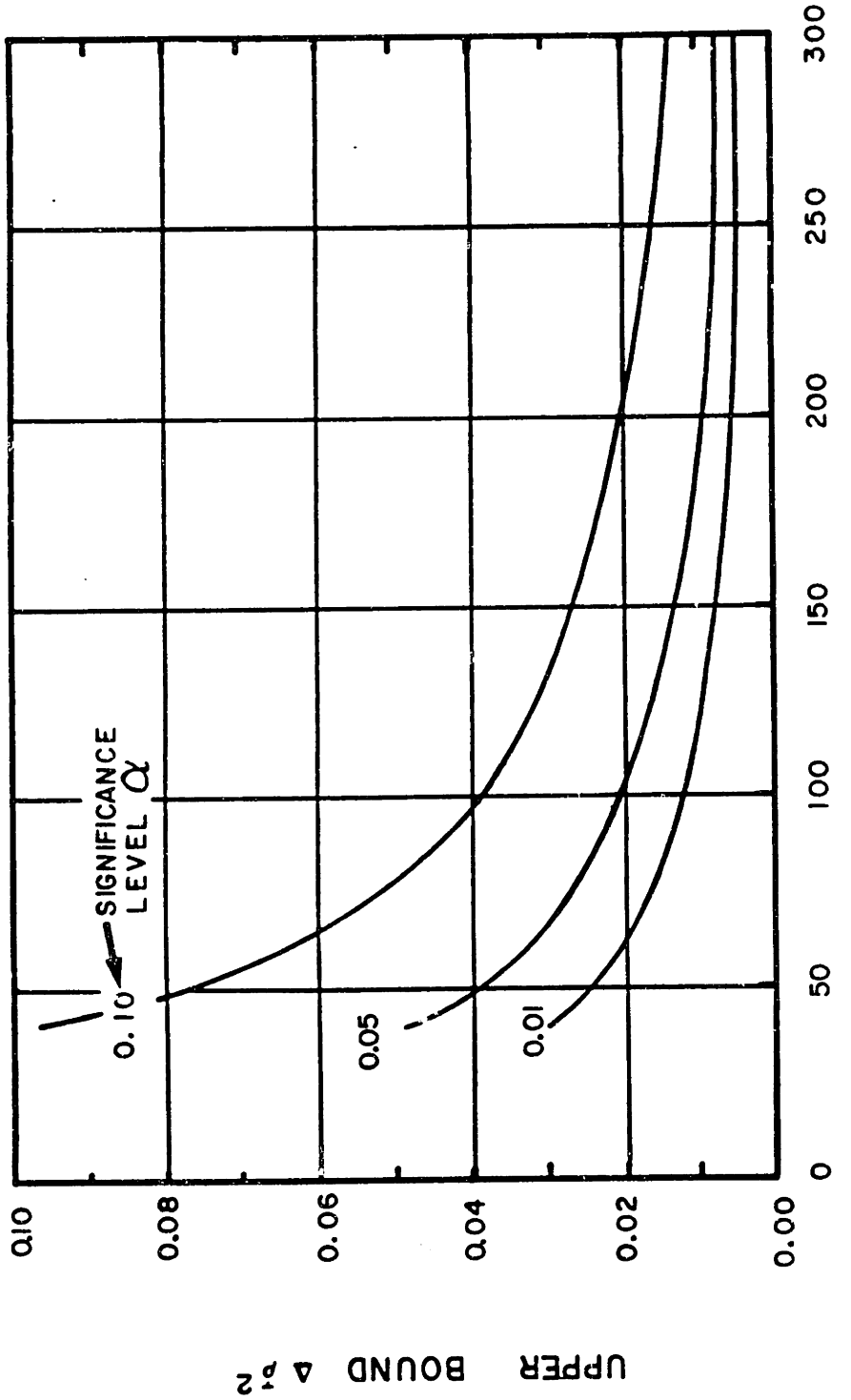
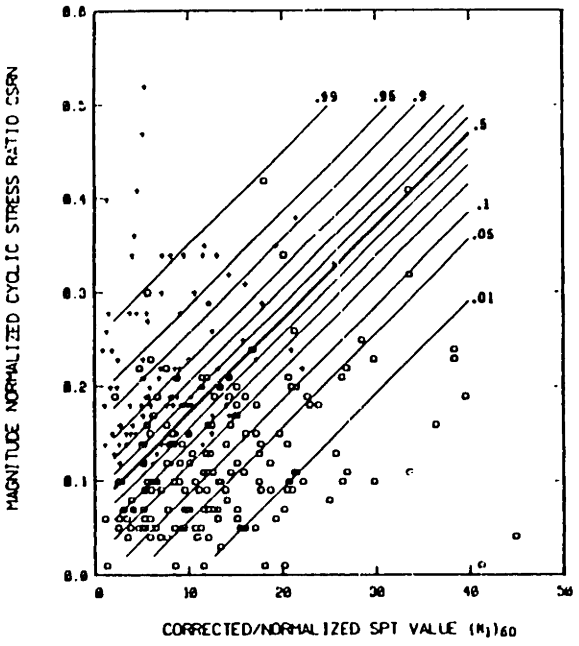
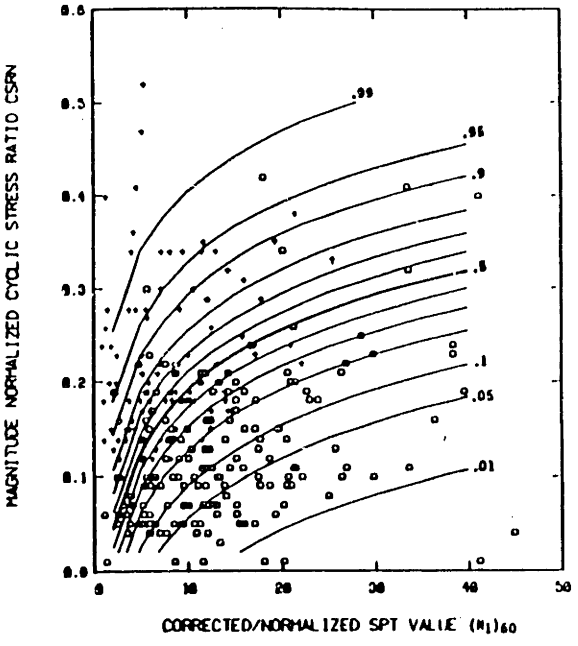


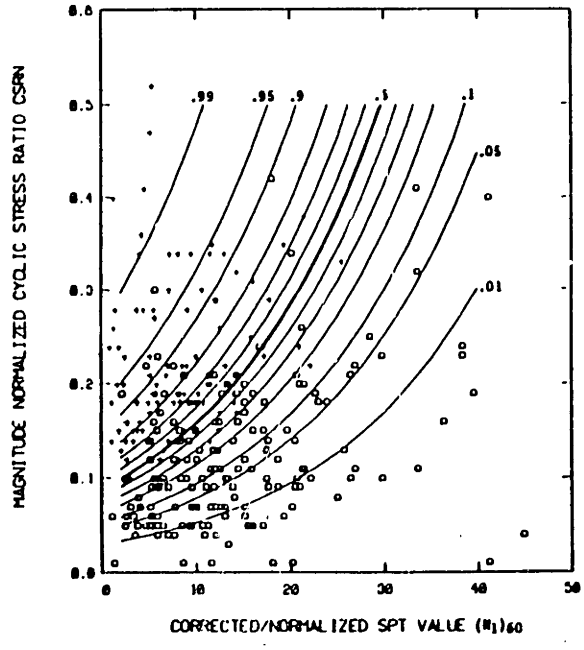
Fig. 4.3 Upper bound values of $\Delta\bar{\rho}^2$ indicating a significant difference in $\bar{\rho}^2$ between 2 models, based on results from Horowitz (1982).



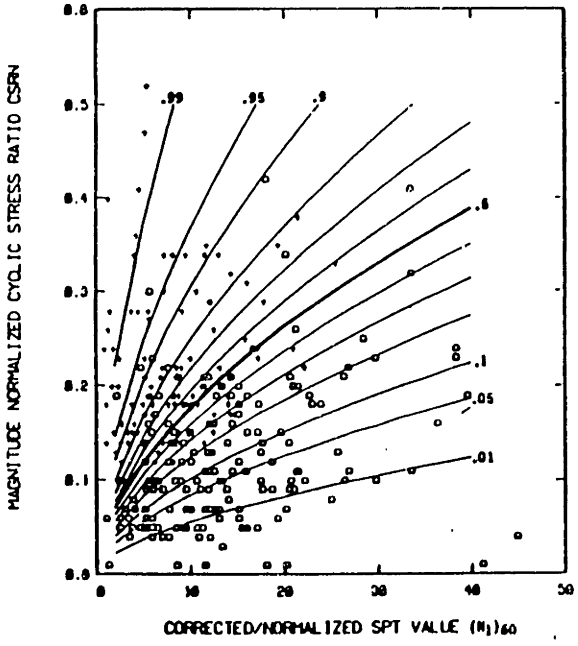
(a) x_1 = CSRN
$x_2 = (N_1)_{60}$



(b) x_1 = CSRN
$x_2 = \ln[(N_1)_{60}]$

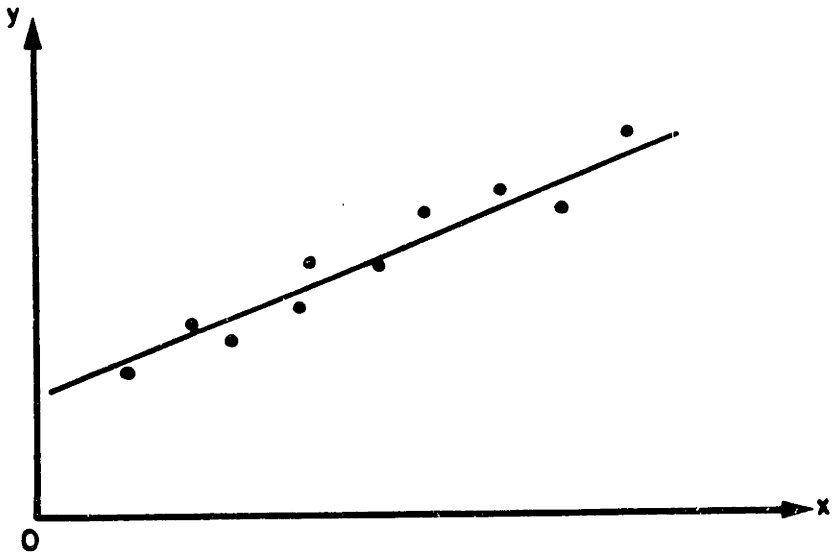


(c) $x_1 = \ln(\text{CSRN})$
$x_2 = (N_1)_{60}$

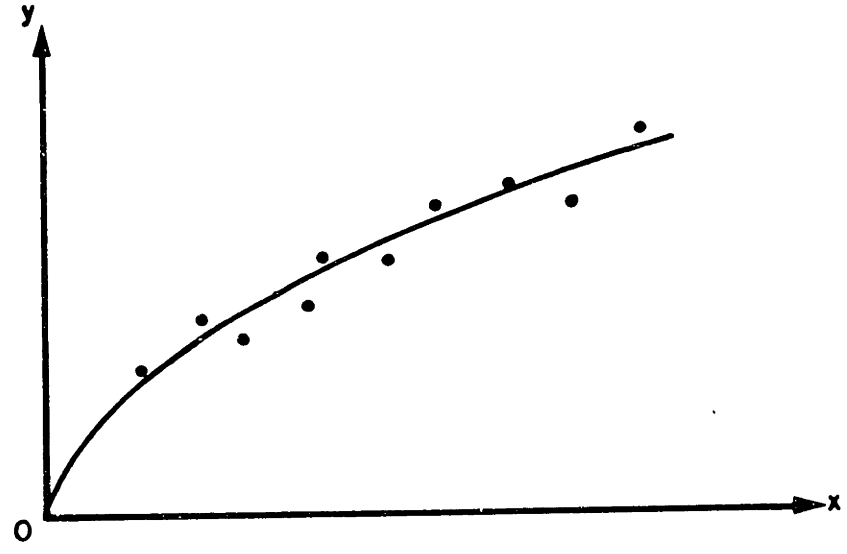


(d) $x_1 = \ln(\text{CSRN})$
$x_2 = \ln[(N_1)_{60}]$

Fig. 4.4 Comparison of shapes of contour lines of equal probability of liquefaction depending on the parametric forms assumed in logit regression.

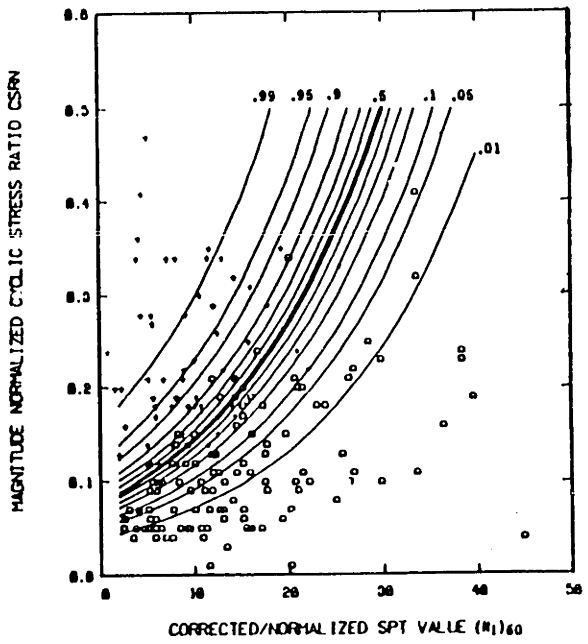


(a) STRAIGHT LINE FITTED TO THE DATA SHOWING A POSITIVE Y-AXIS INTERCEPT

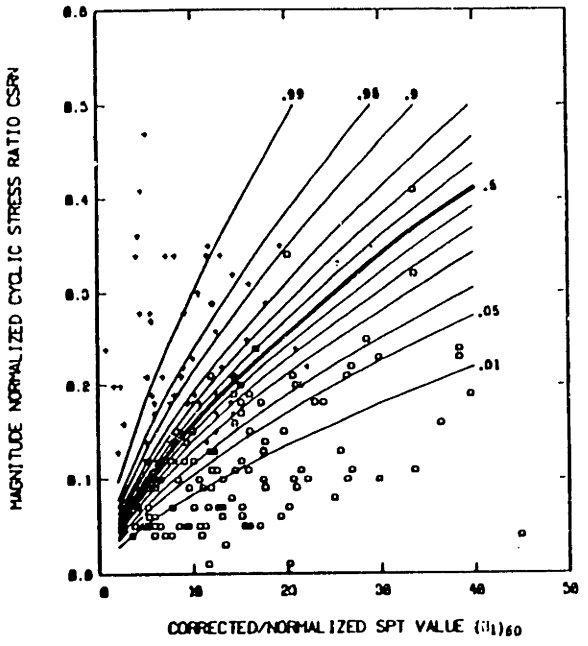


(b) POLYNOMIAL CURVE OR POWER FUNCTION CONSTRAINED TO PASS THROUGH THE ORIGIN FITTED TO THE SAME DATA AS IN (a)

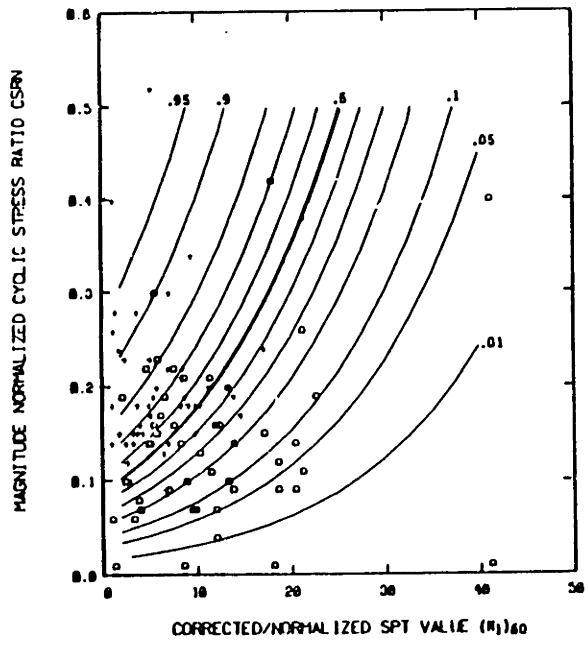
Fig. 4.5 Schematic of a explanation for the concave downward curvature of a fitted line constrained to pass through the origin.



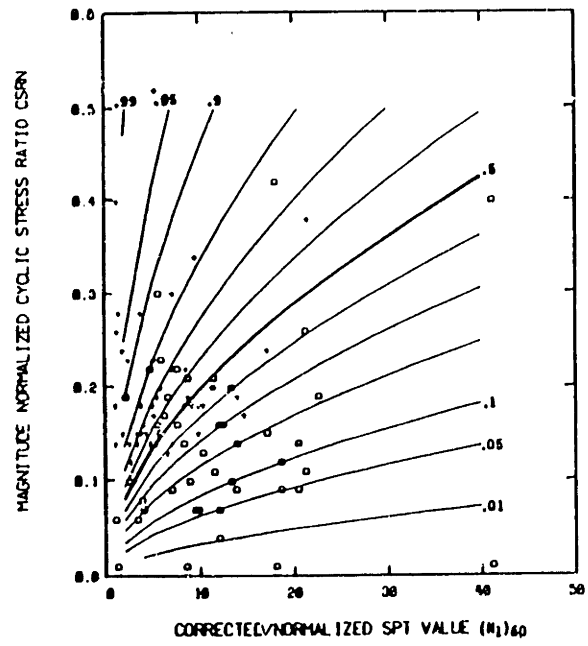
(a) $x_1 = \ln(CSRN)$
$x_2 = (N_1)_{60}$
Clean sand



(b) $x_1 = \ln(CSRN)$
$x_2 = \ln[(N_1)_{60}]$
Clean sand

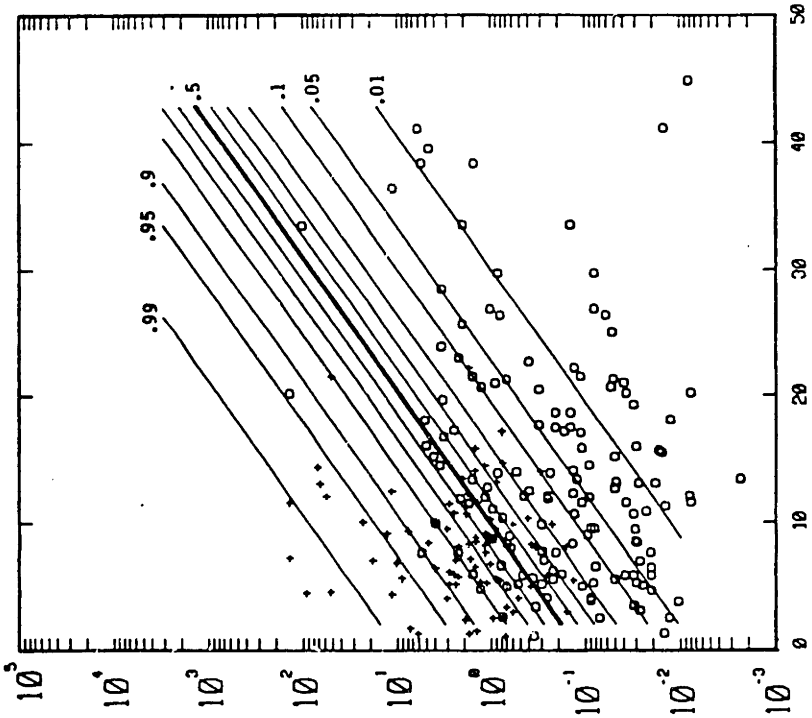


(c) $x_1 = \ln(CSRN)$
$x_2 = (N_1)_{60}$
Silty sand

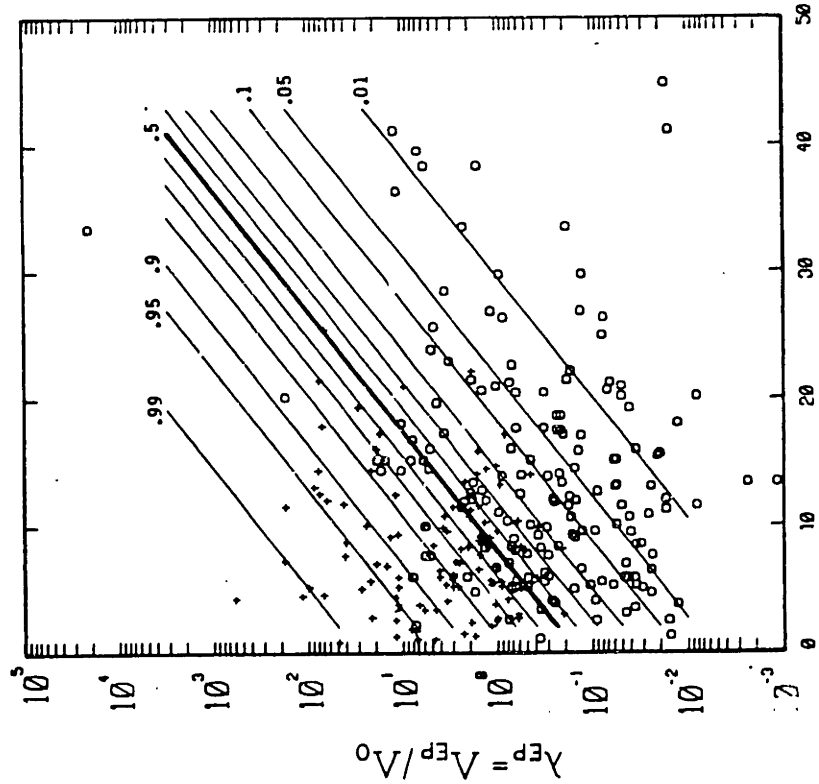


(d) $x_1 = \ln(CSRN)$
$x_2 = \ln[(N_1)_{60}]$
Silty sand

Fig. 4.6 Comparison of shapes of contour lines of equal probability of liquefaction depending on the parametric forms assumed in logit regression for clean sand and silty sand cases.



(b) Hypocentral (modified) model based on 228 case subset with focal depth data.



(a) Epicentral model based on all 278 data cases.

Fig. 4.7 Contours of equal probability of liquefaction for Davis-Berrill and modified Davis-Berrill models.

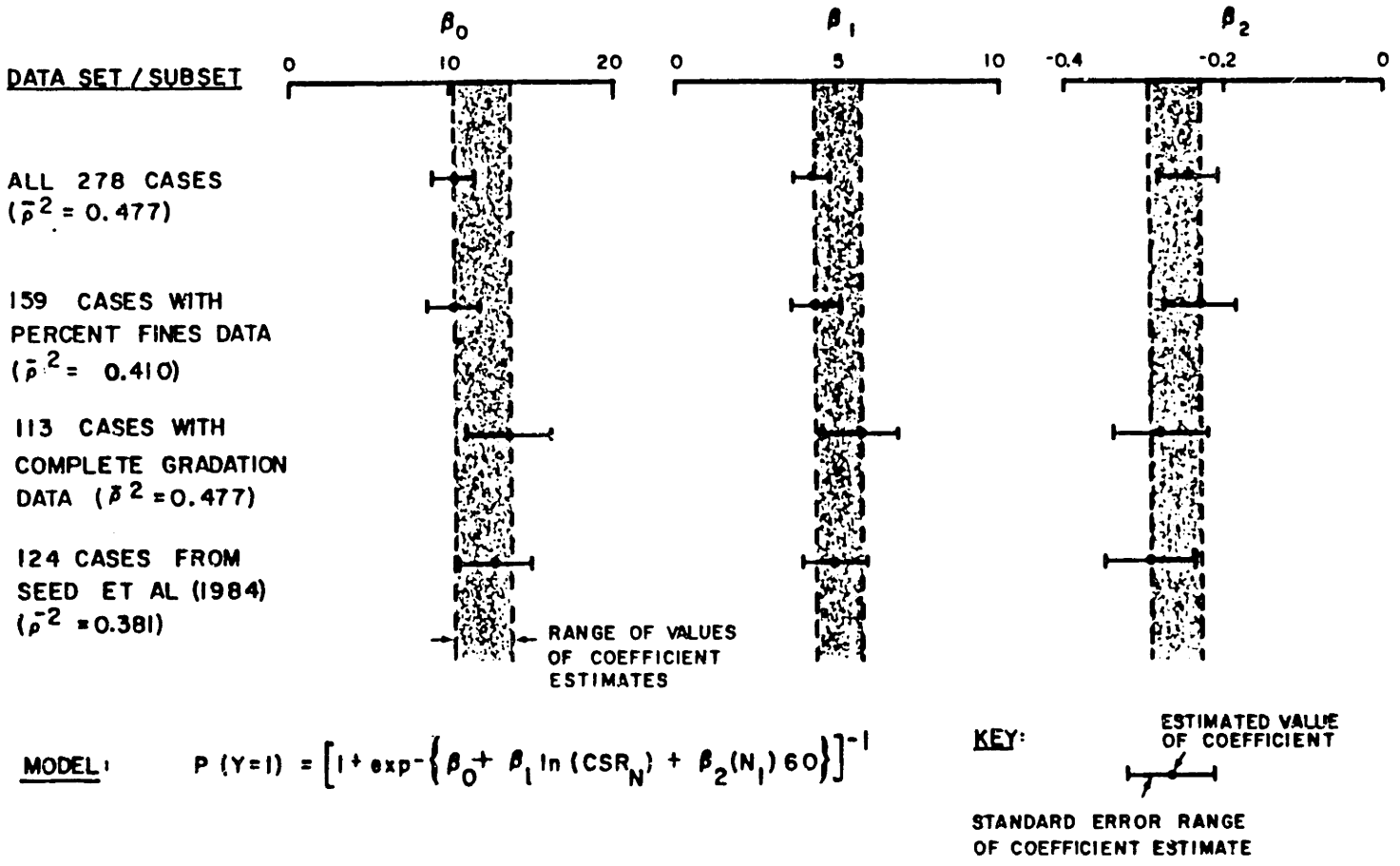
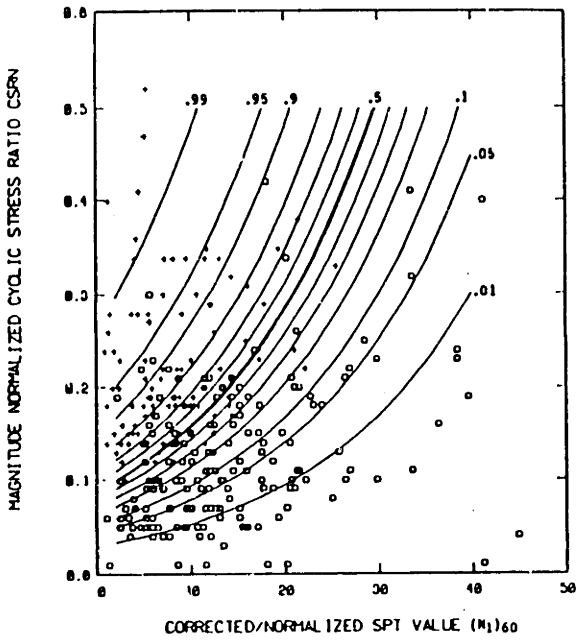
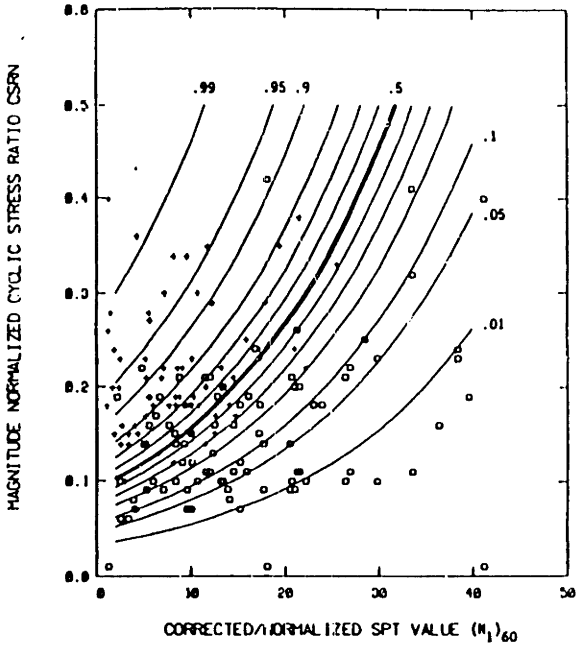


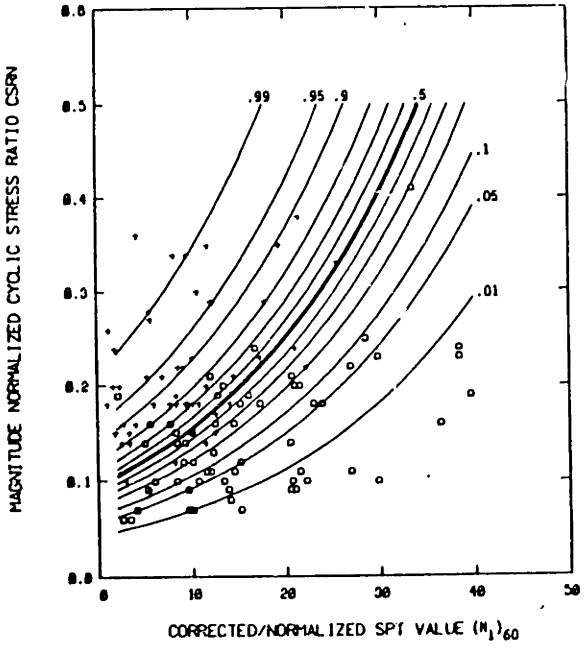
Fig. 4.8 Comparison of logit coefficient estimates for various data sets.(No partitioning into silty and clean sand subsets.)



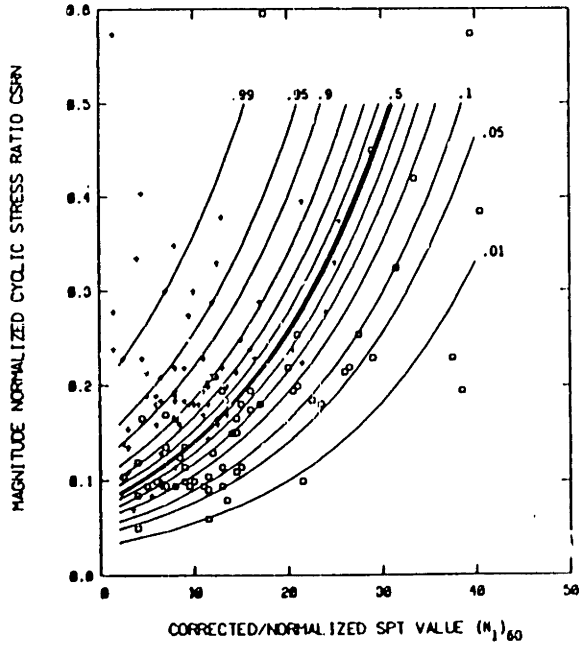
(a) All 278 cases



(b) 159 cases with fines content data



(c) 113 cases with all gradation data



(d) 124 cases from Seed et al.

Fig. 4.9 Contours of equal probability of liquefaction fitted for various data sets or subsets. (No partitioning of data into clean and silty sands.)

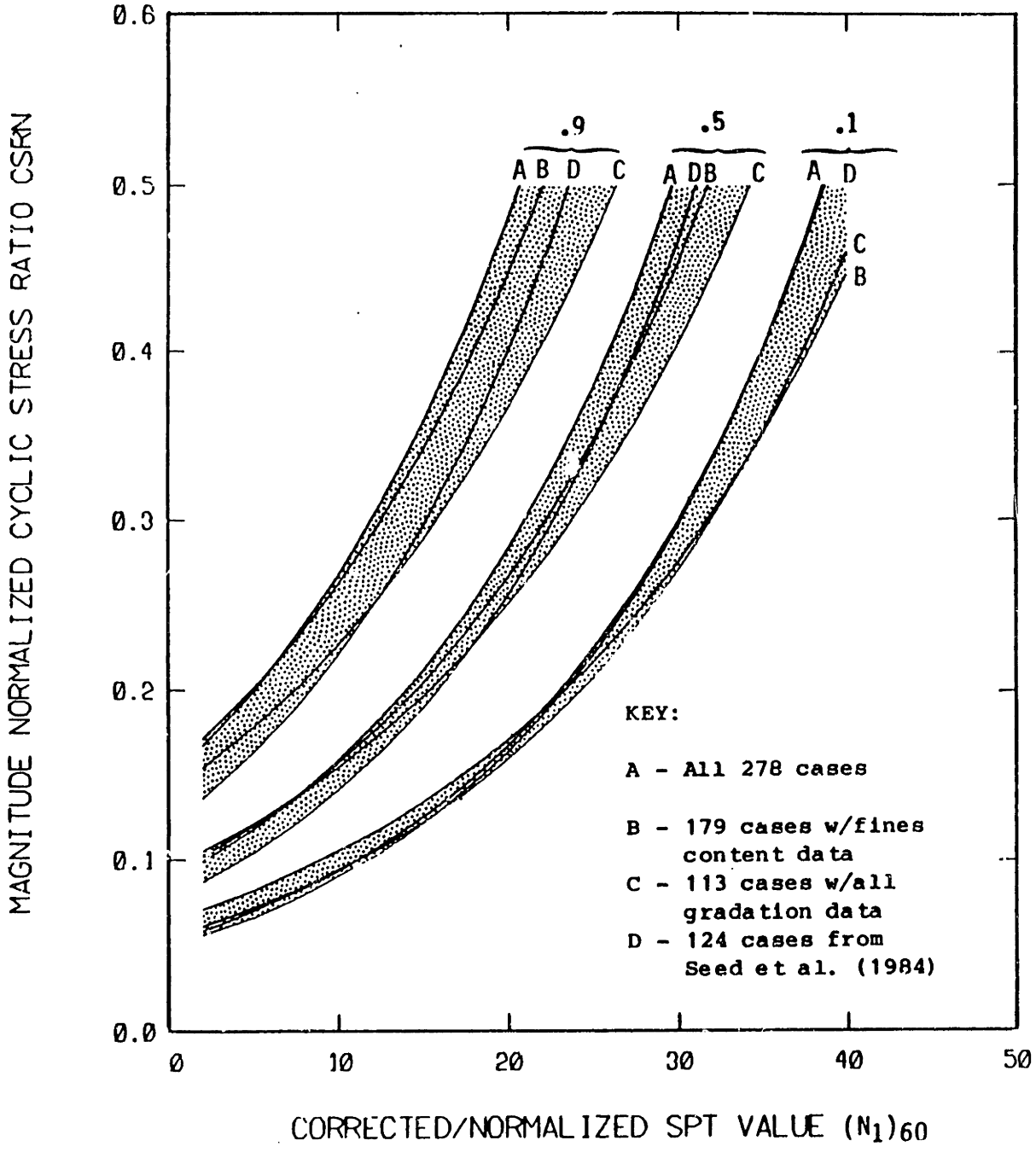


Fig. 4.10 Comparison of the 0.1, 0.5, and 0.9 countours of equal probability of liquefaction fitted to various data sets or subsets. (No partitioning of data into clean and silty sands.)

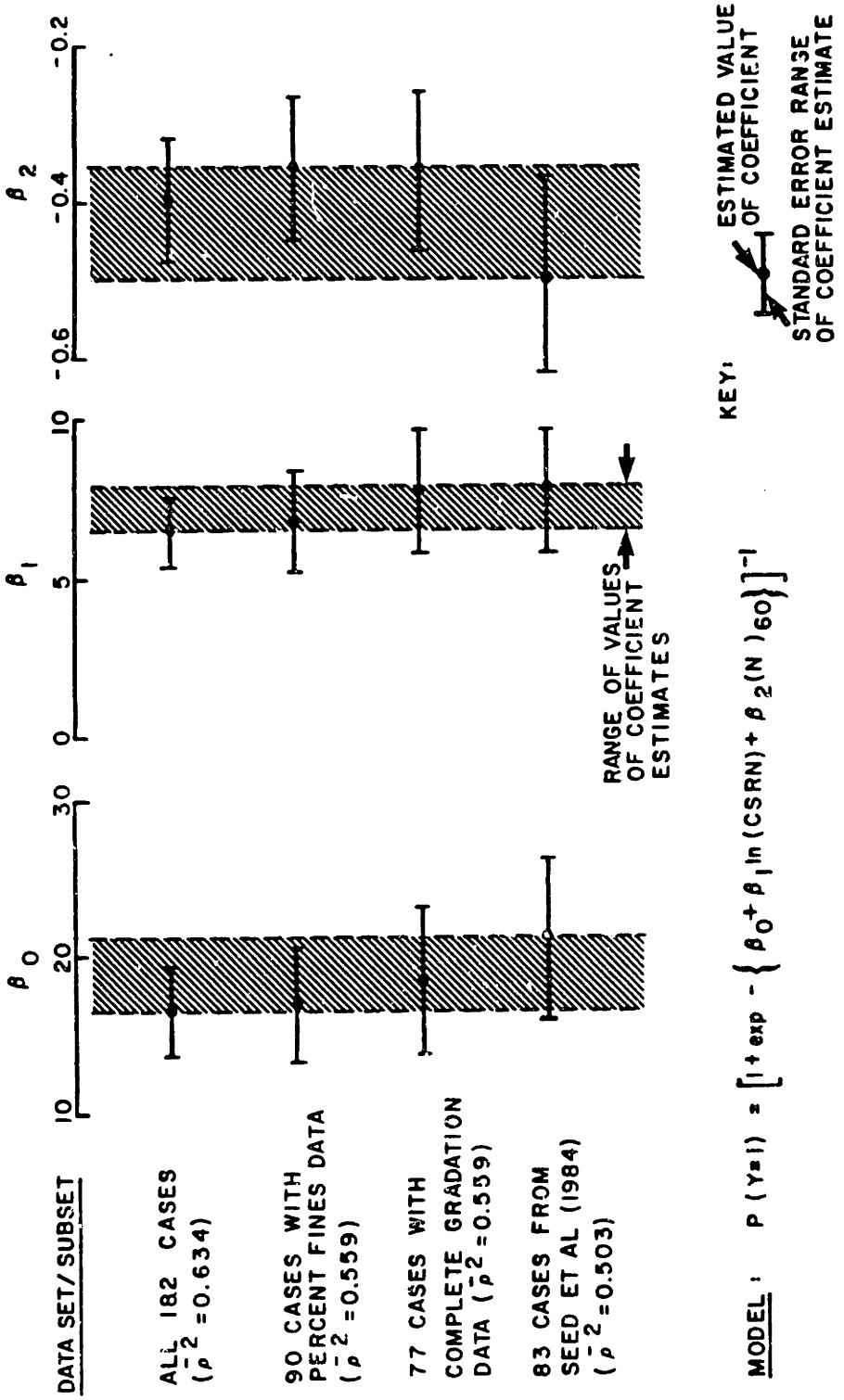


Fig. 4.11 Comparison of logit coefficient estimates for various data sets. Clean sand data.

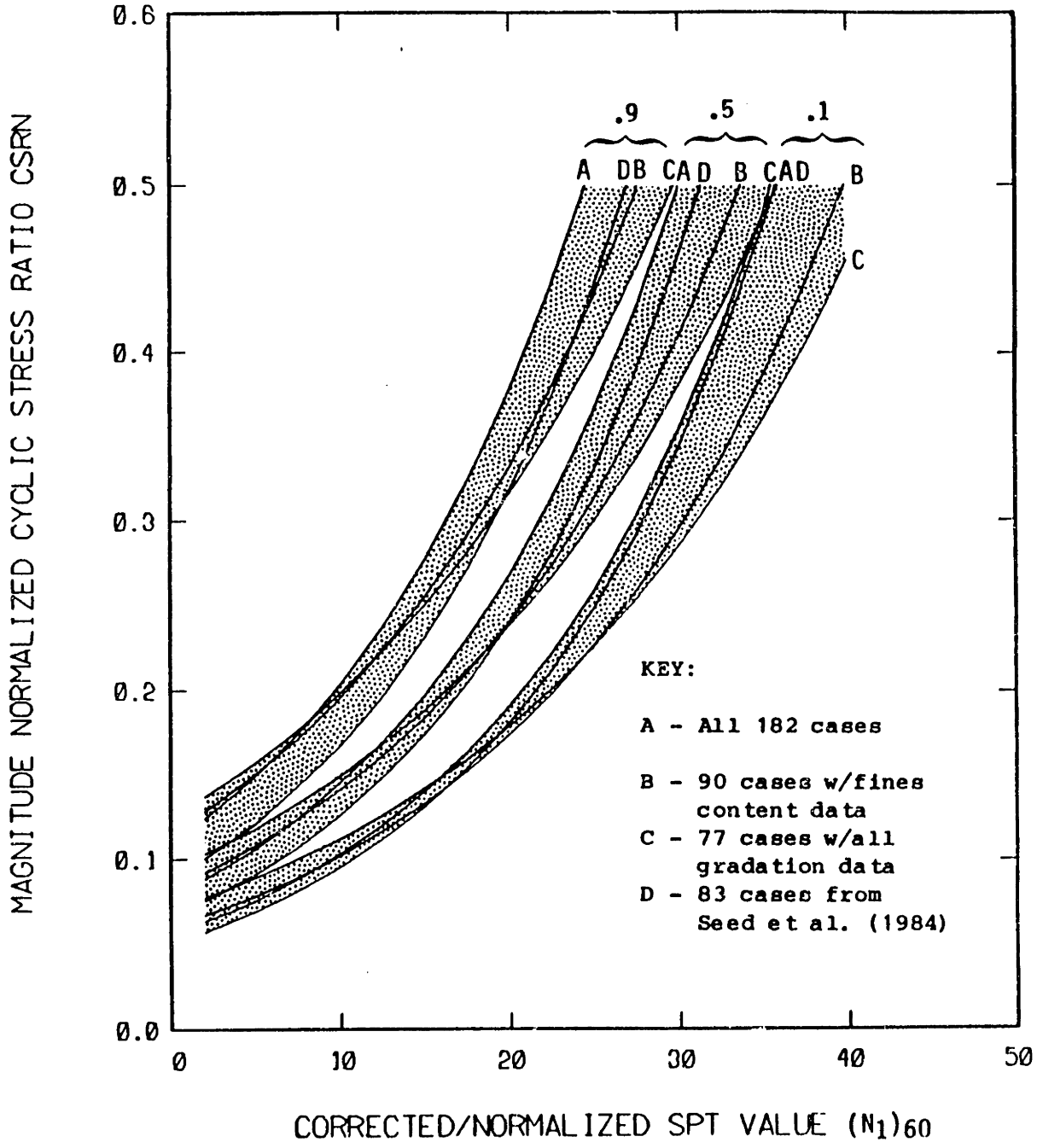


Fig. 4.12 Comparison of the 0.1, 0.5, and 0.9 contours of equal probability of liquefaction fitted to various data sets or subsets. Clean sand data (FC < 12%).

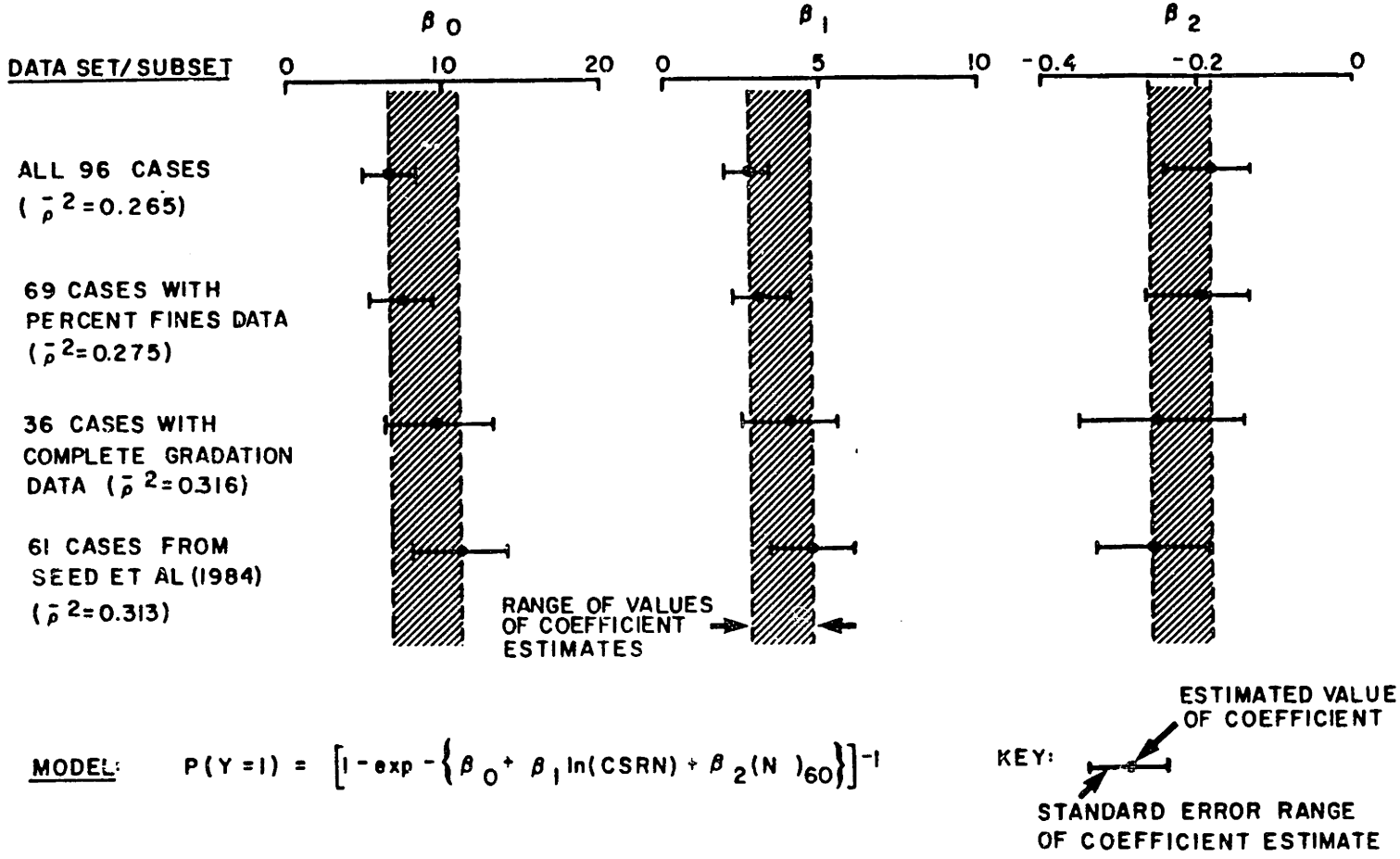


Fig. 4.13 Comparison of logit coefficient estimates for various data sets. Silty sand data (FC > 12%).

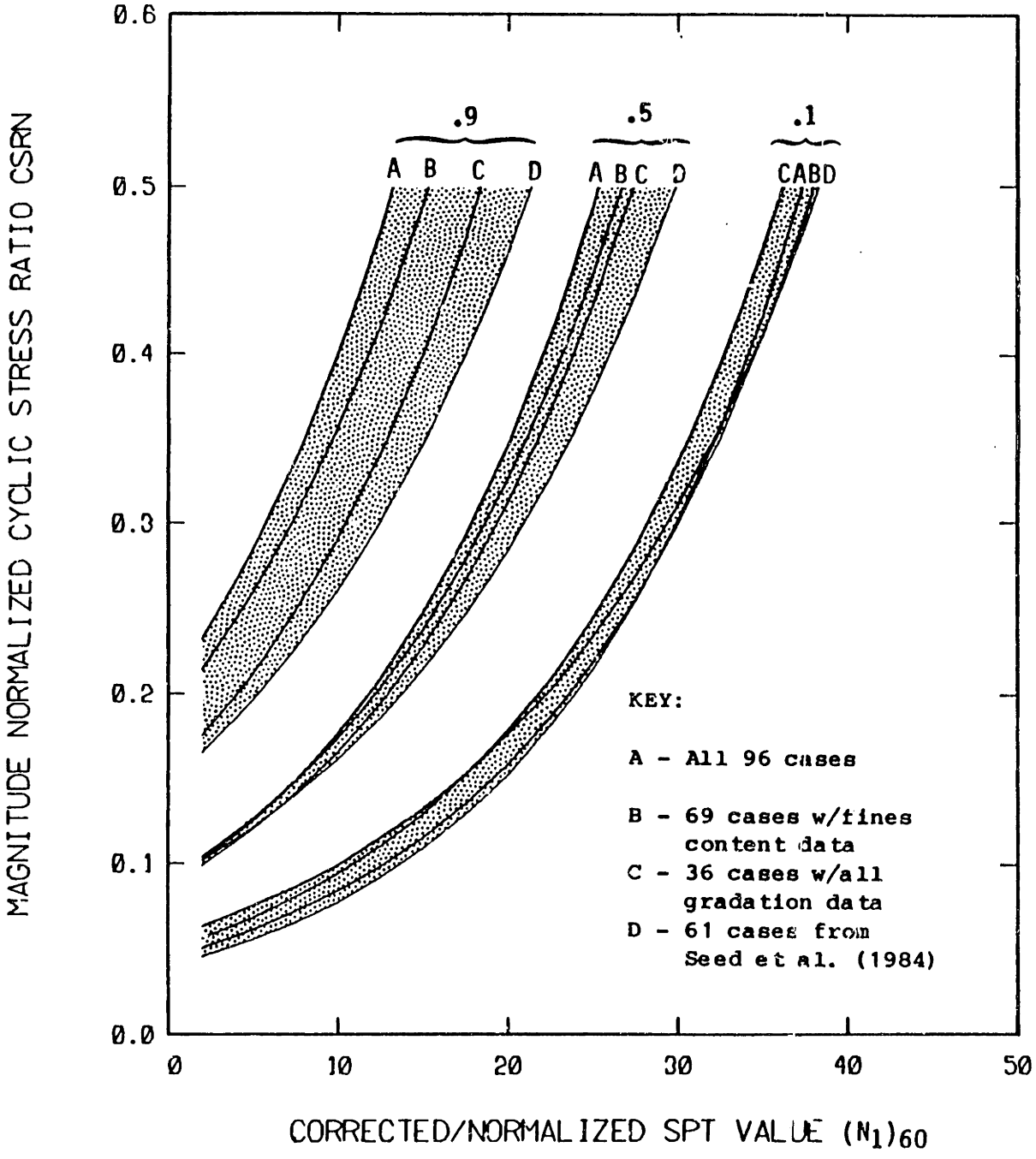


Fig. 4.14 Comparison of the 0.1, 0.5 and 0.9 contour of equal probability of liquefaction fitted to various data sets or subsets. Silty sand data (FC > 12%).

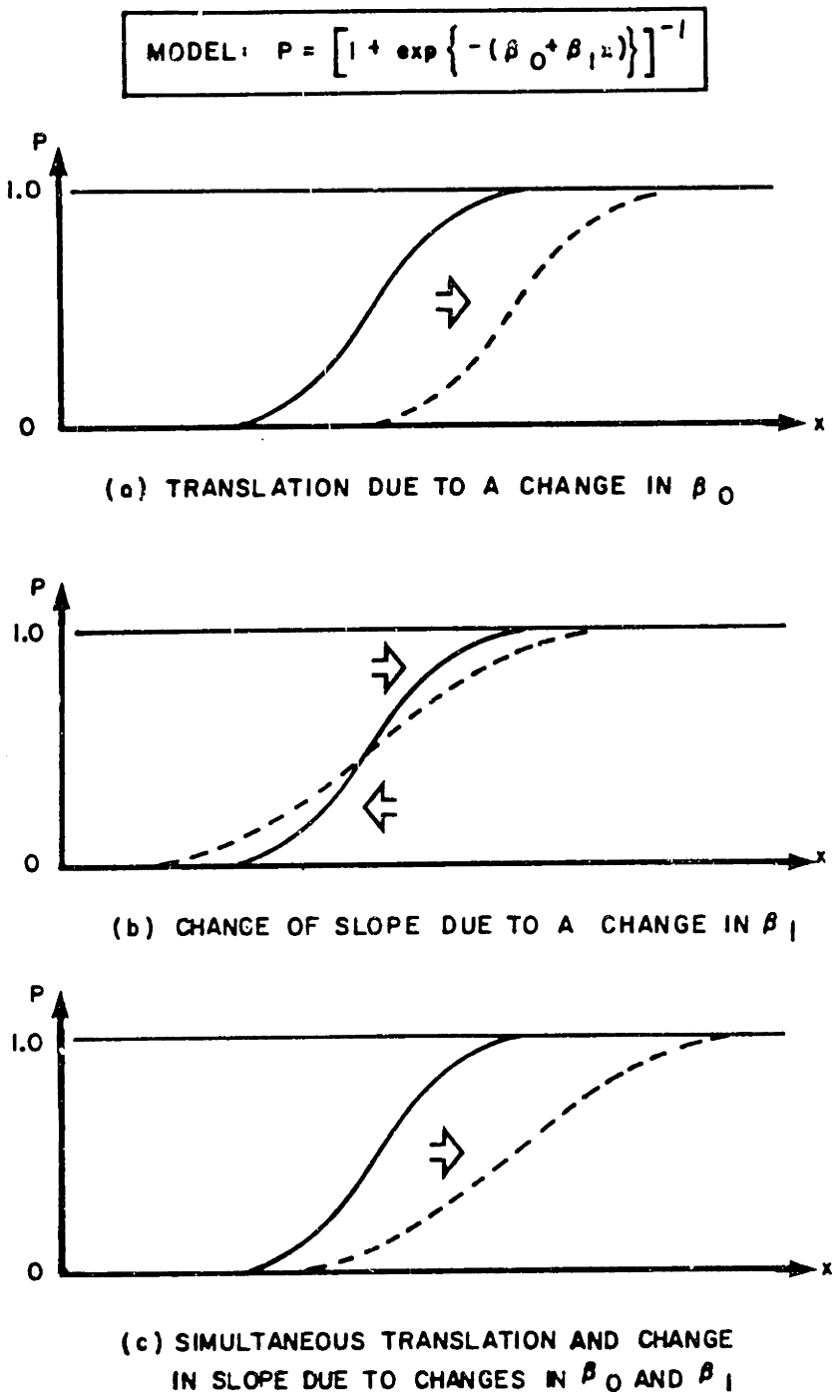


Fig. 4.15 Schematic of the interpretation of changes in the logit regression coefficients β for a model with a single explanatory variable x.

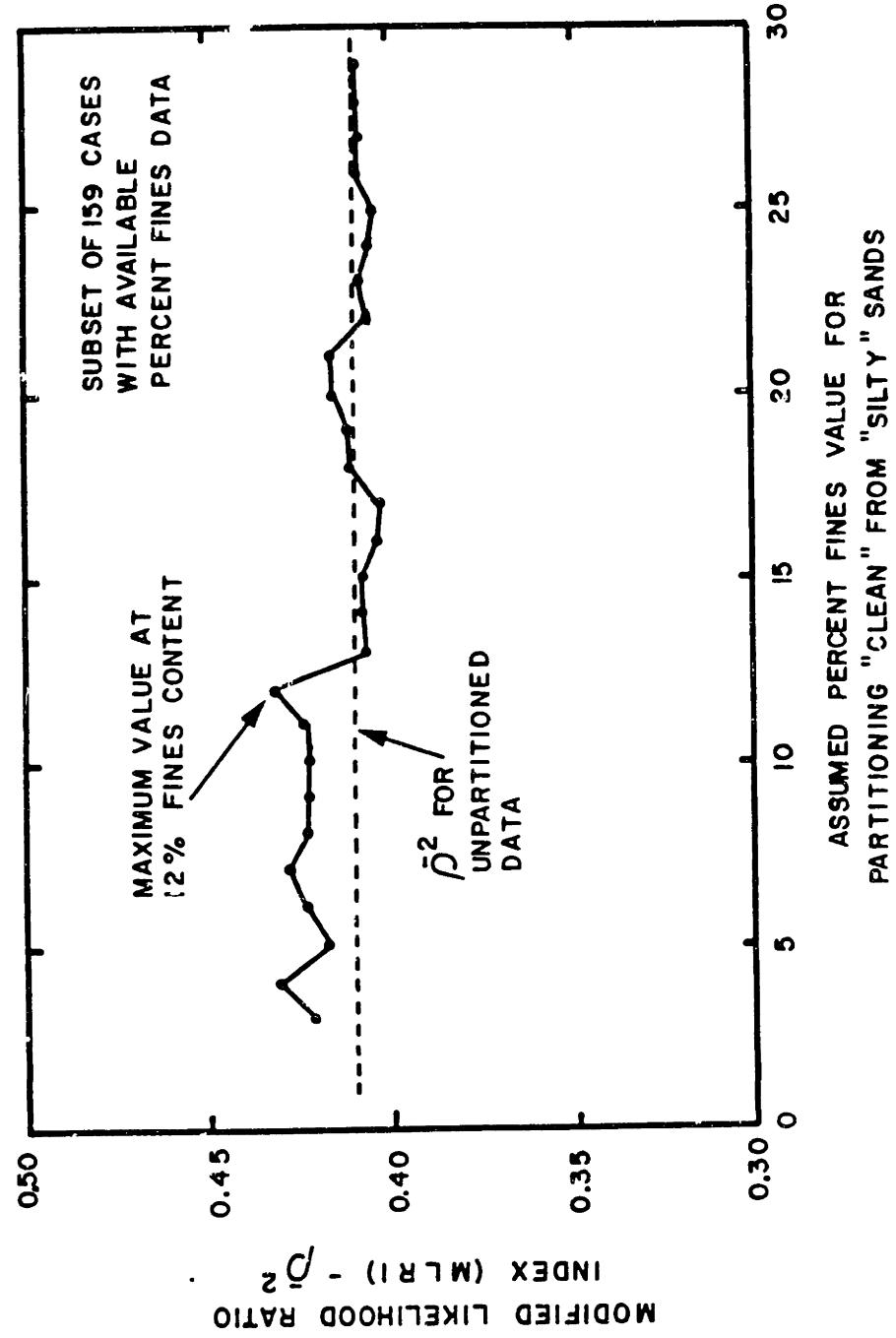


Fig. 4.16 Modified likelihood ratio index plotted against the assumed percent fines value for partitioning "clean" from silty sands - 159 case subset with FC data.

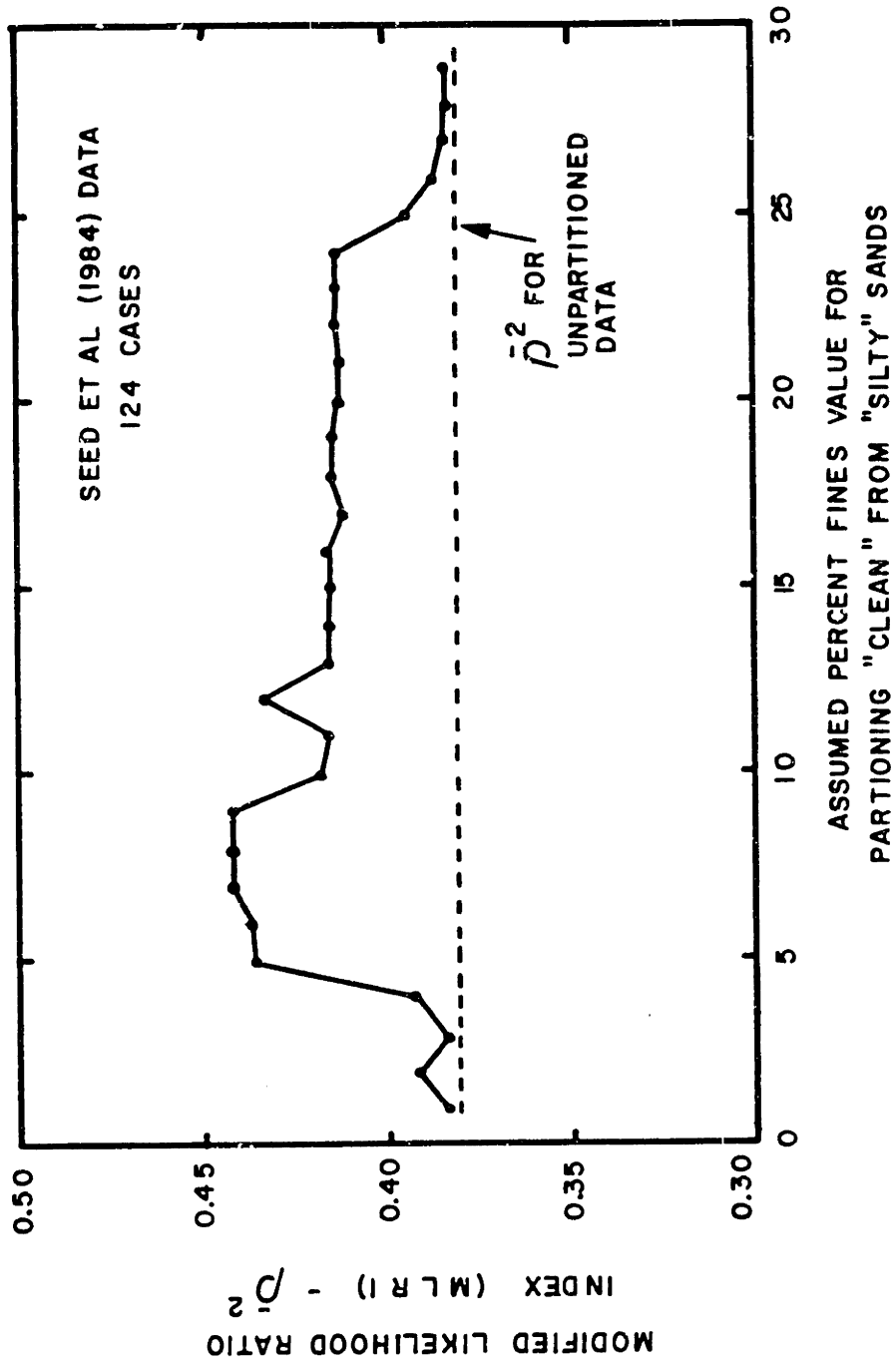


Fig. 4.17 Modified likelihood ratio index plotted against the assumed percent fines value for partitioning "clean" from silty sands - Seed et al. (1984) data set.

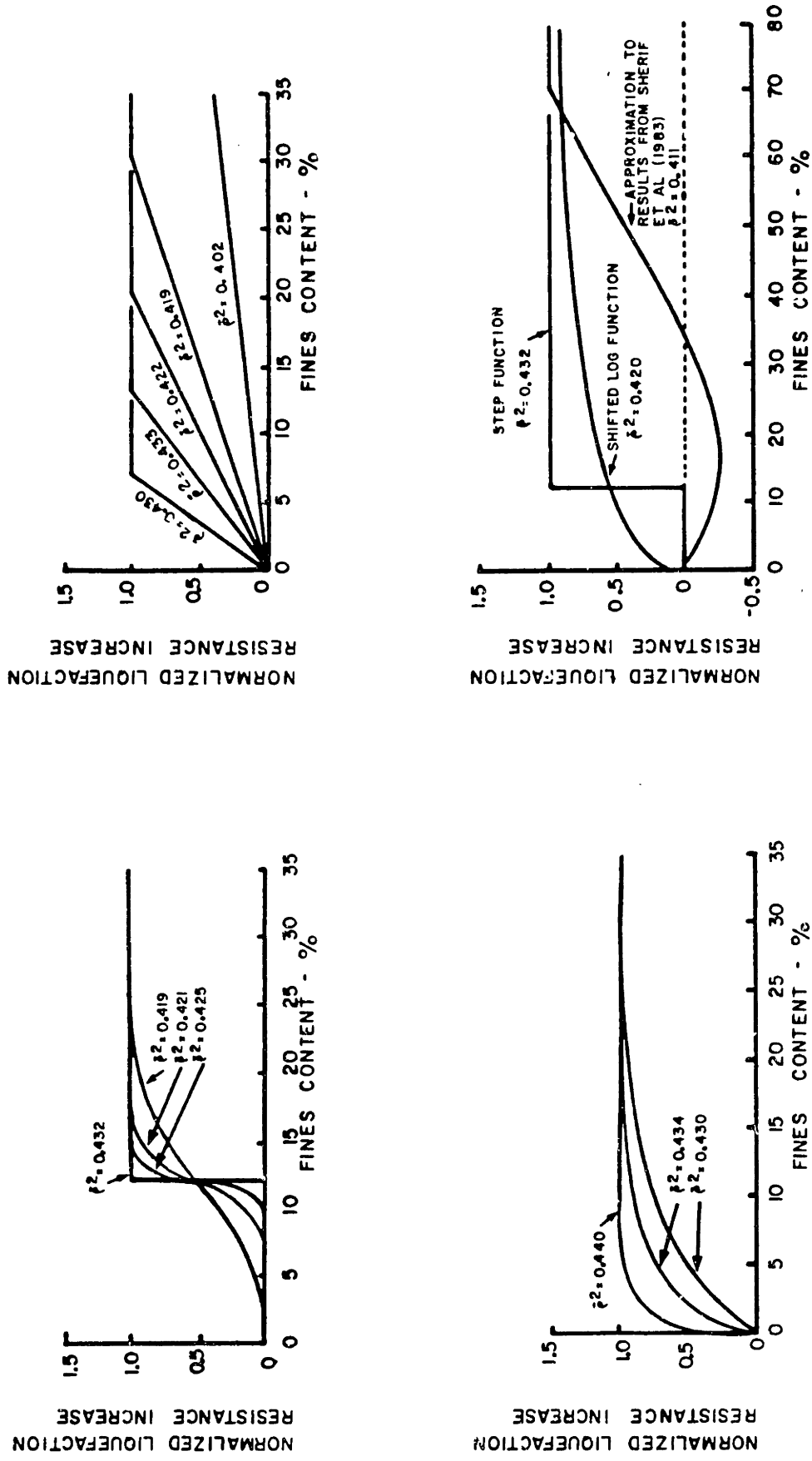
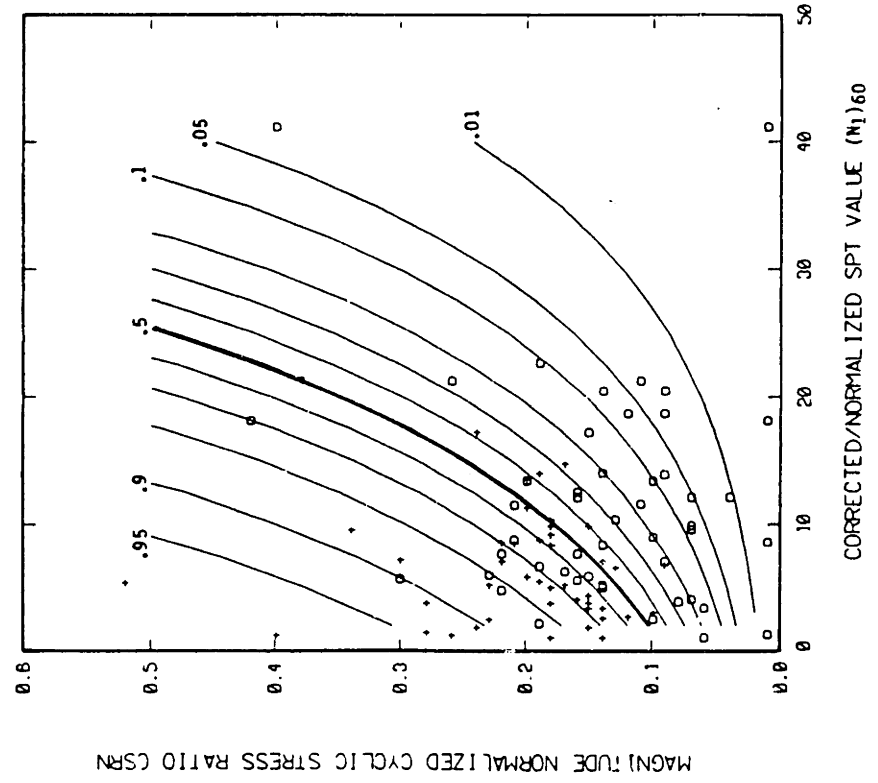
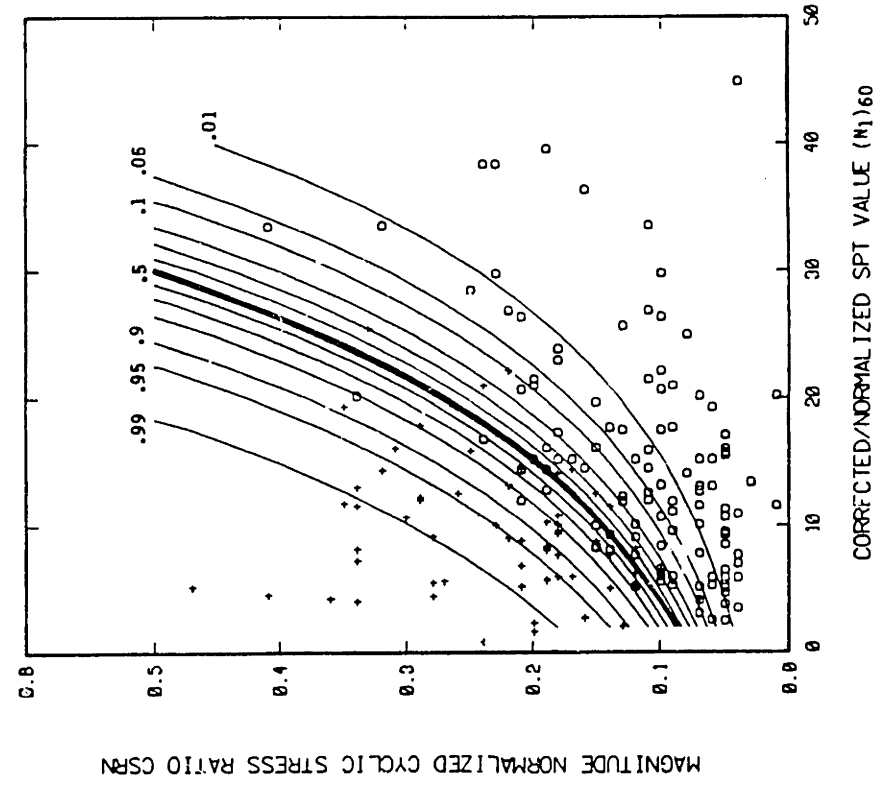


Fig. 4.18 Comparison of $\bar{\rho}^2$ values for various types of functions assumed for modeling the normalized liquefaction resistance increase as a function of fines content.



(b) Silty sand (FC > 12%)



(a) Clean Sand (FC < 12%)

Fig. 4.19 Contours of equal probability of liquefaction for the clean and silty sand for Seed-Idriss model considering fines content only.

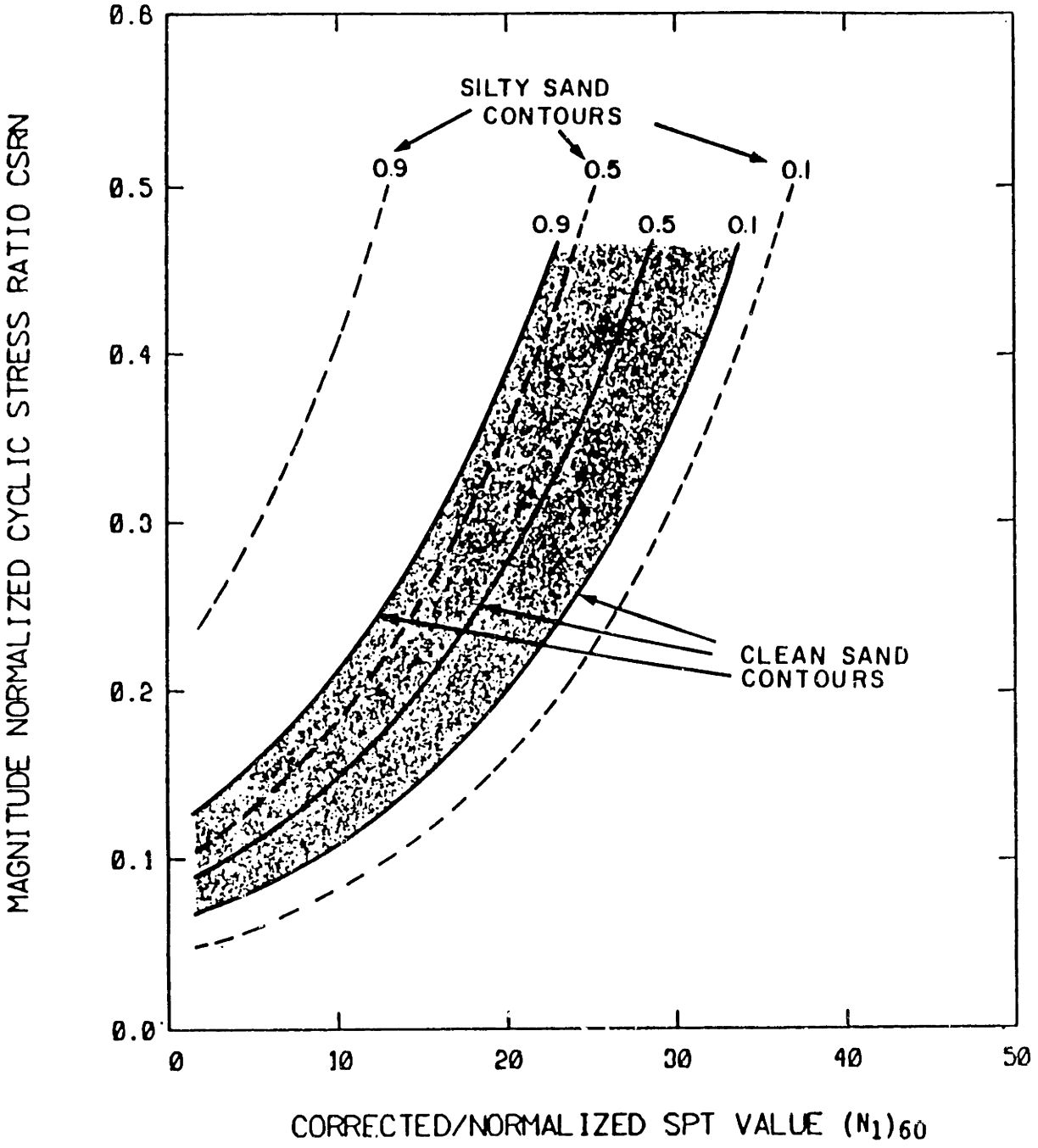


Fig. 4.20 Comparison of 0.1, 0.5, and 0.9 contours of equal probability of liquefaction for Seed-Idriss model considering fines content only.

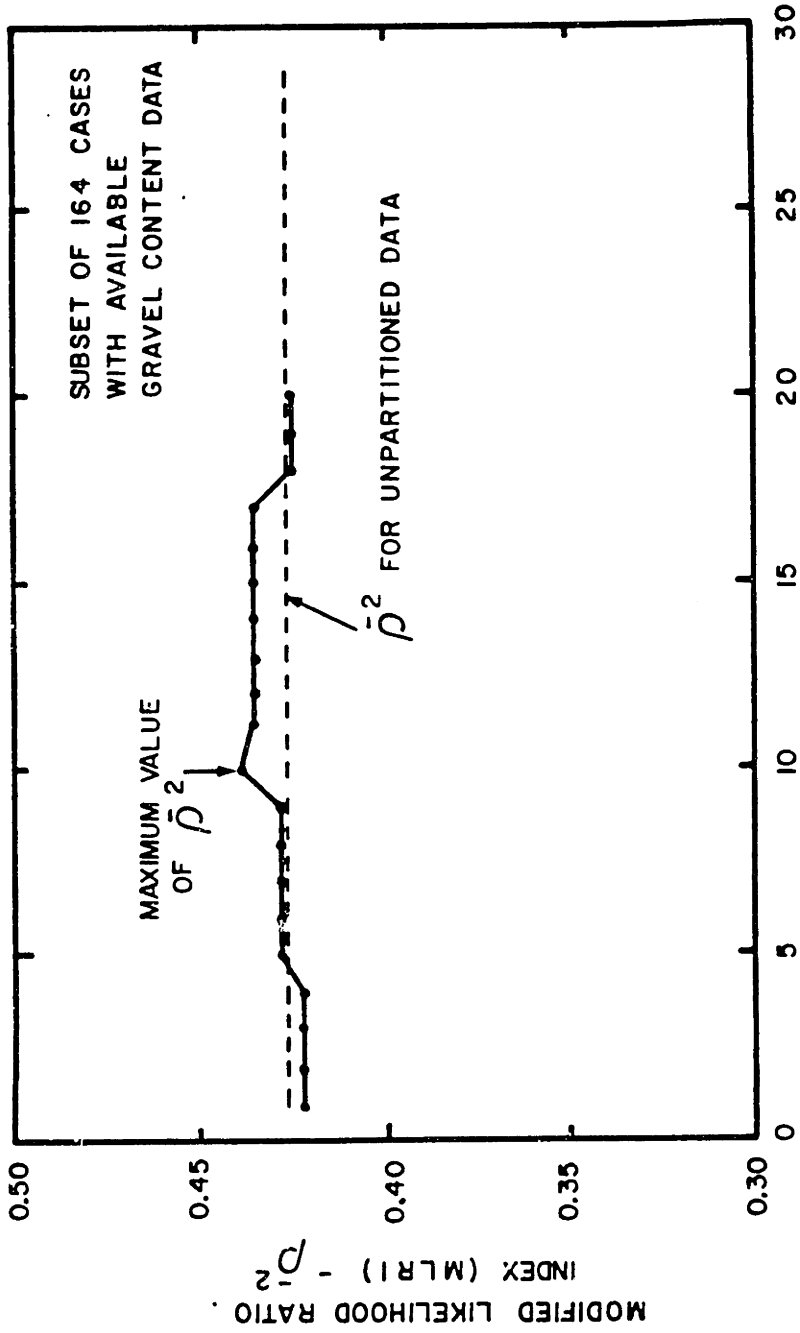
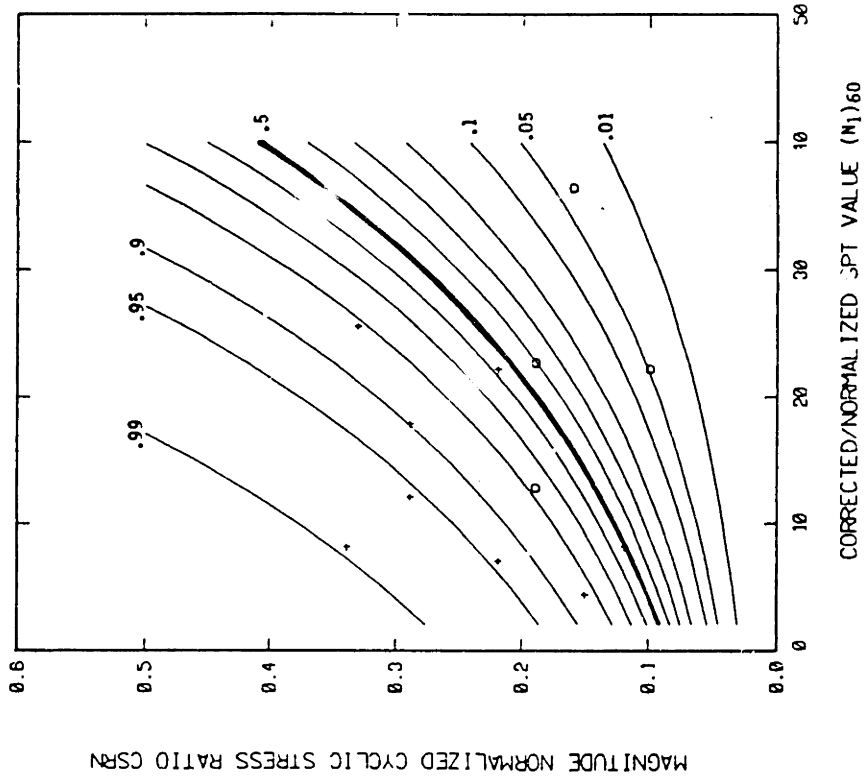


Fig. 4.21 Modified likelihood ratio index plotted against the assumed percent gravel value partitioning "gravelly" from "non-gravelly" sands - 164 case subset with GC data.



(a) Non-gravelly sand (GC < 10%).

(b) Gravelly sand (GC > 10%).

Fig. 4.22 Contours of equal probability of liquefaction for gravelly and non-gravelly sand for Seed-Idriss model considering gravel content only.

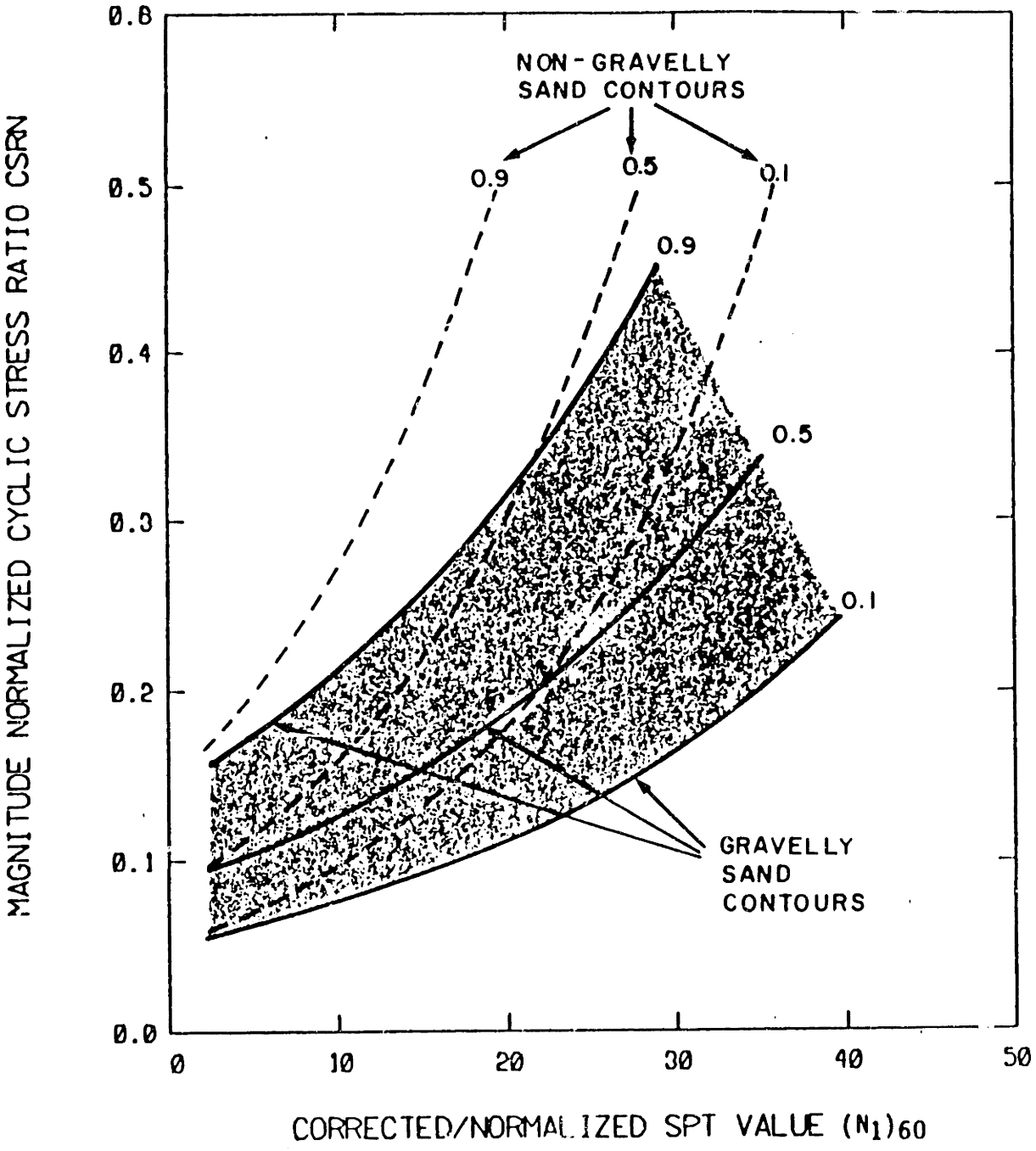
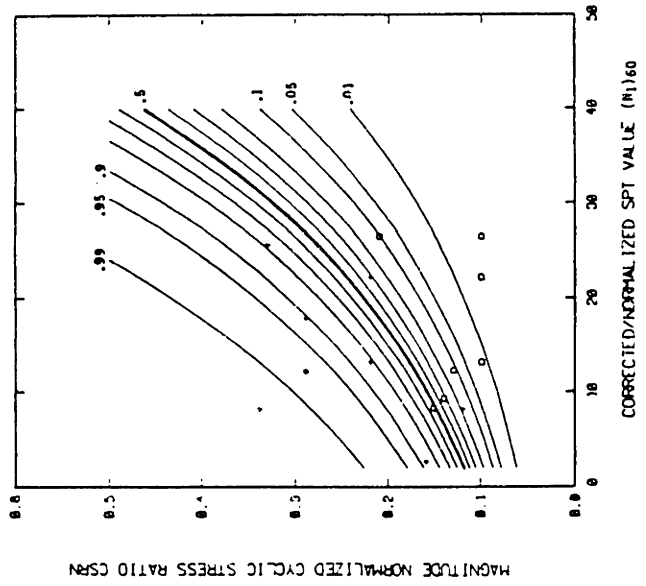
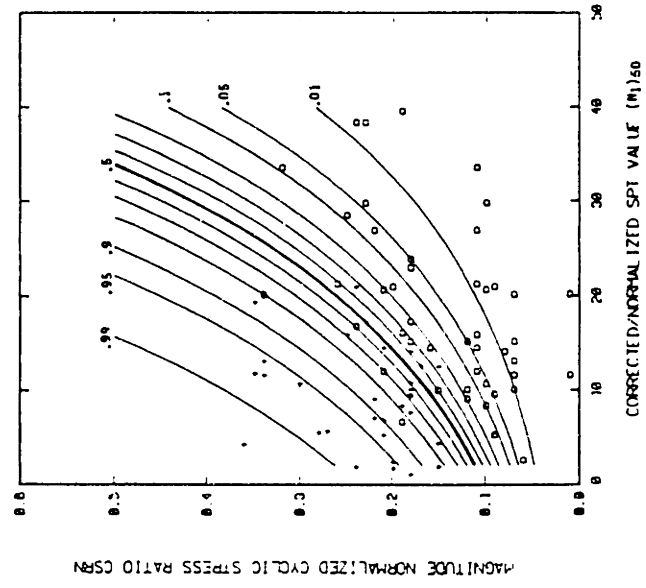


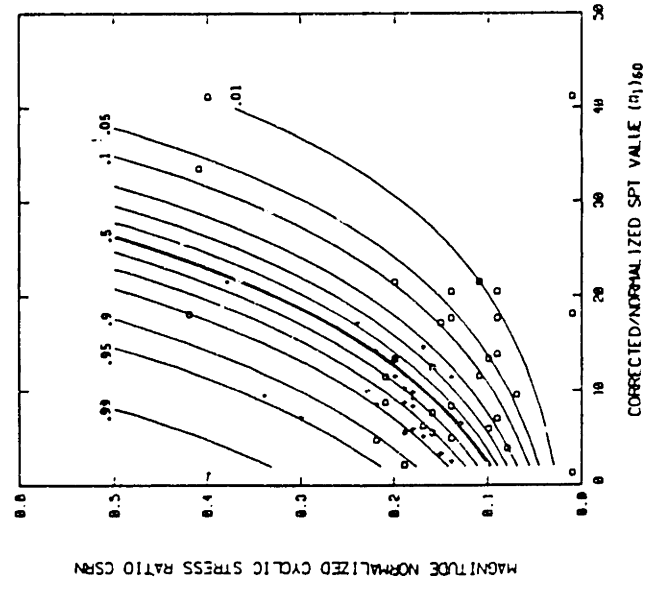
Fig. 4.23 Comparison of P = 0.1, 0.5, and 0.9 contours of equal probability of liquefaction for Seed-Idriss model considering gravel content only.



(c) D_{50} = 1.0
(Data plotted for
$0.6 \leq D_{50} \leq 2.0$)

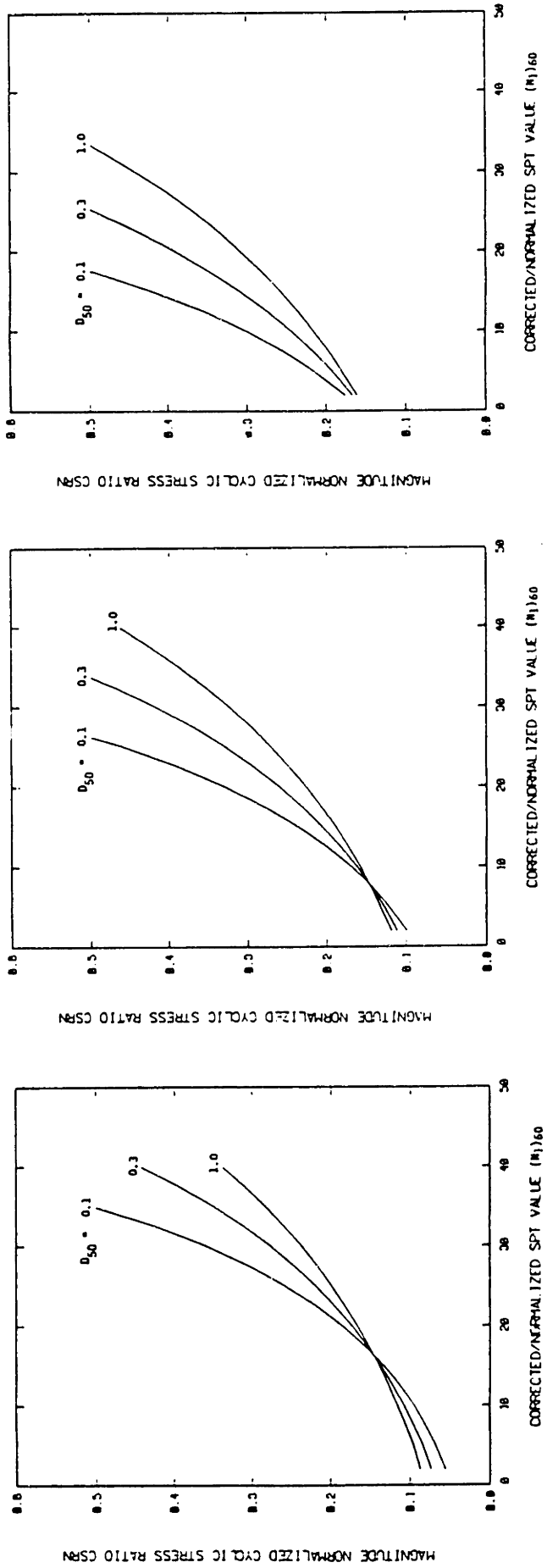


(b) D_{50} = 0.3
(Data plotted for
$0.2 \leq D_{50} \leq 0.6$)



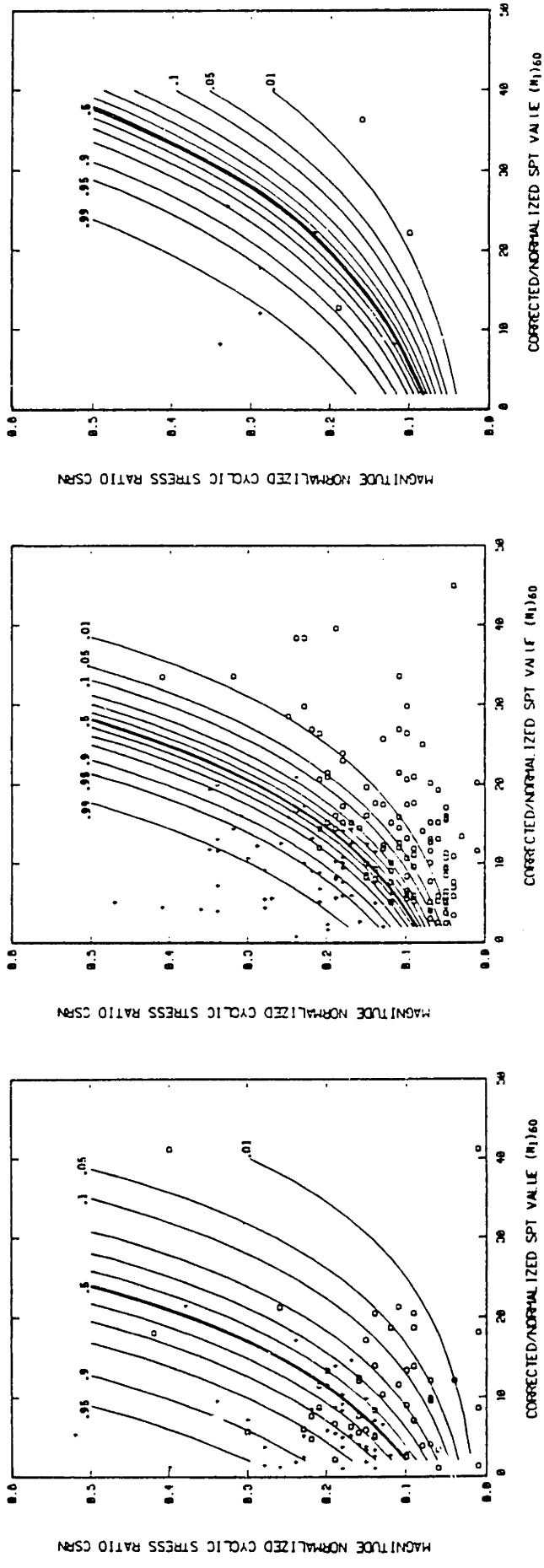
(a) D_{50} = 0.1
(Data plotted for
$0.06 \leq D_{50} \leq 0.2$)

Fig. 4.24 Contours of equal probability of liquefaction for various values of median grain size D_{50} for Seed-Idriss model considering D_{50} effects only.



(a) P = 0.1 (b) P = 0.5 (c) P = 0.9

Fig. 4.25 Comparison of P = 0.1, 0.5, and 0.9 contours of equal probability of liquefaction for Seed-Idriss model considering various values of median grain size D_{50}.

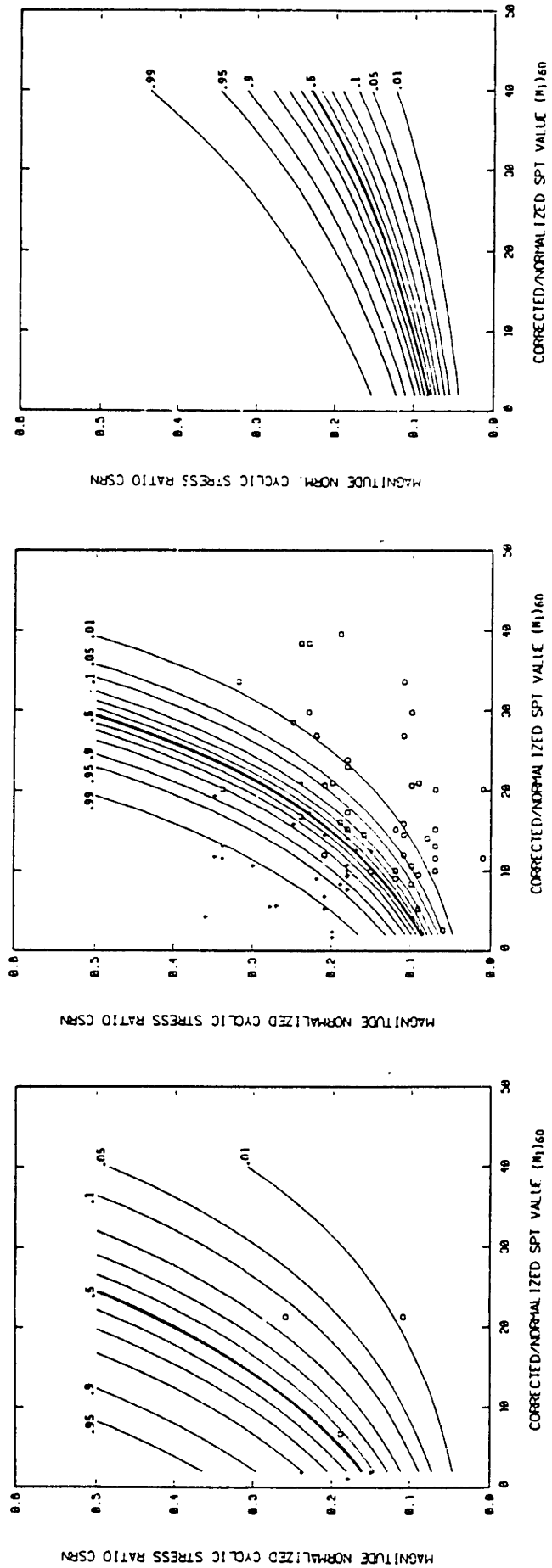


(a) Silty sand
FC > 12%
GC < 10%

(b) Clean sand
FC < 12%
GC < 10%

(c) Gravelly sand
FC < 10%
GC > 10%

Fig. 4.26 Comparison of contours of equal probability of liquefaction for Seed-Idriss model incorporating effects of fines content and gravel content (fitted using all 278 cases).



(a) Silty sand
FC > 12%
GC < 10%

(b) Clean sand
FC < 12%
GC < 10%

(c) Gravelly sand
FC < 10%
GC > 10%

Fig. 4.27 Comparison of contours of equal probability of liquefaction for Seed-Idriss model incorporating effects of D_{50} size, fines content and gravel content (fitted using 165 subset of cases with D_{50} size). Contours shown for D_{50} = 0.30 mm. Data plotted are for $0.2 \leq D_{50} < 0.6$.

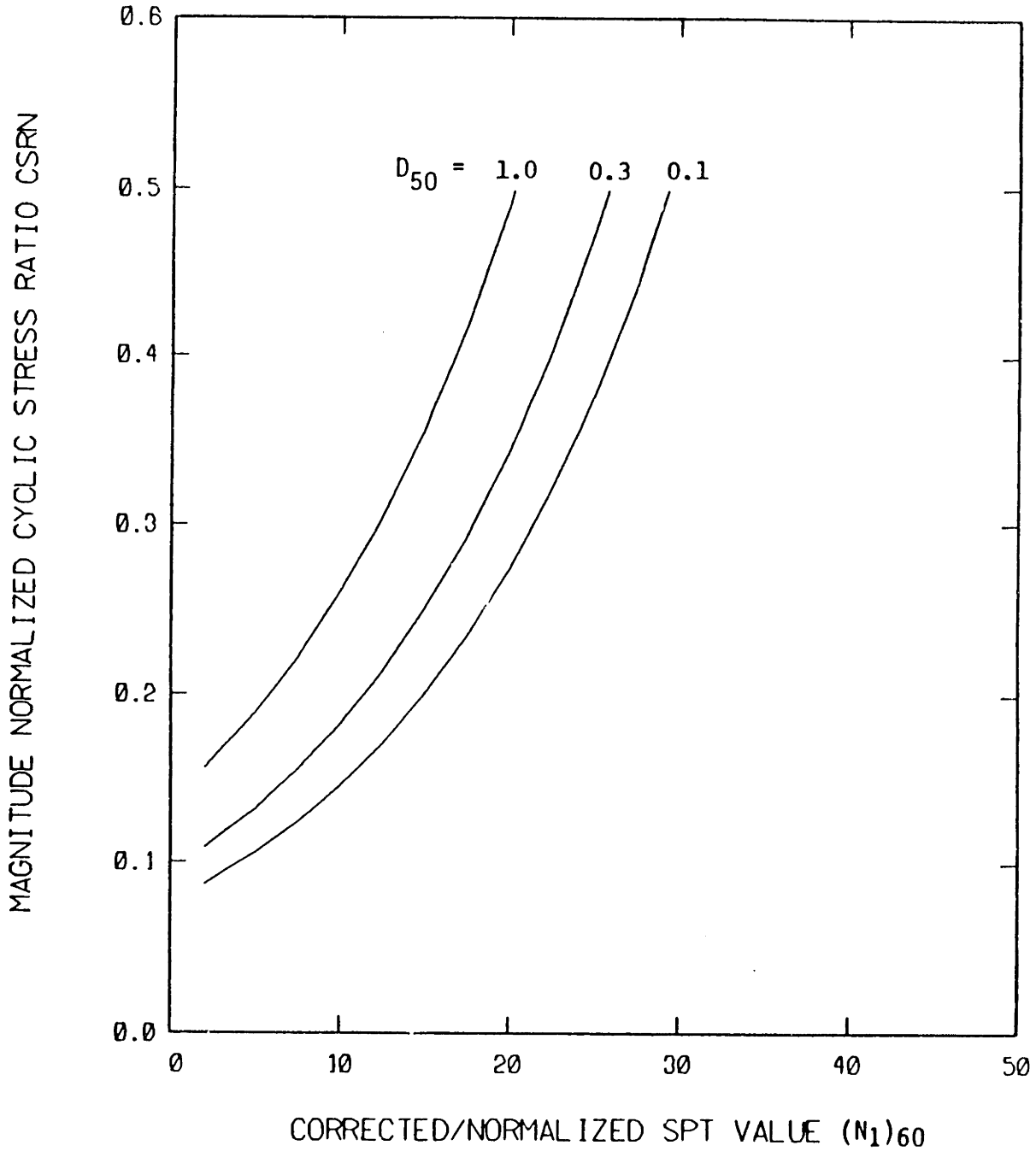
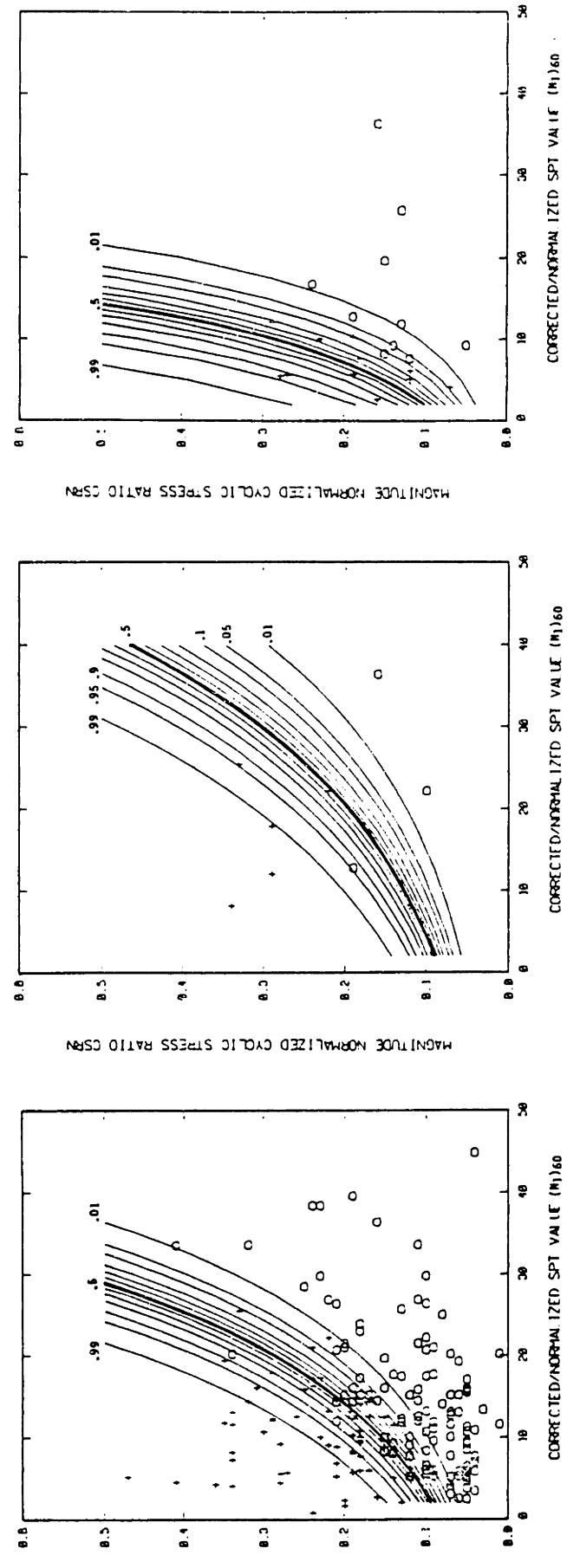


Fig. 4.28 Comparison of P = 0.5 contours of equal probability of liquefaction for Seed-Idress model incorporating effects of D_{50} size, fines content, and gravel content. Contours shown for clean sand with FC < 12% and GC < 10%.



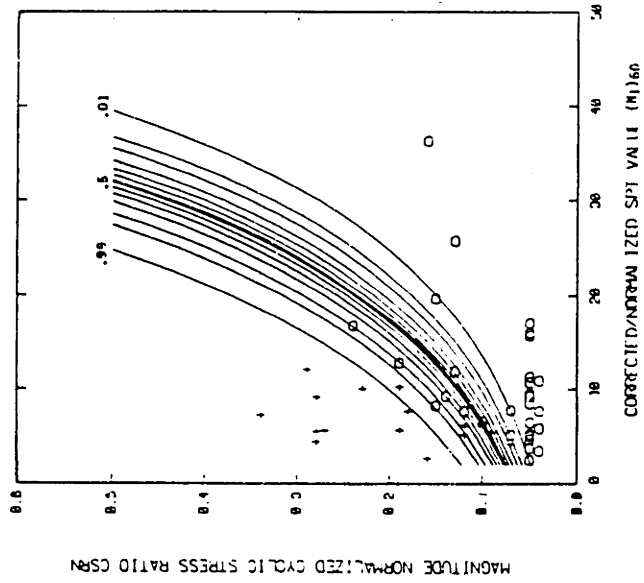
(a) Clean sand "base" model

(b) Gravel Content
GC < 10%

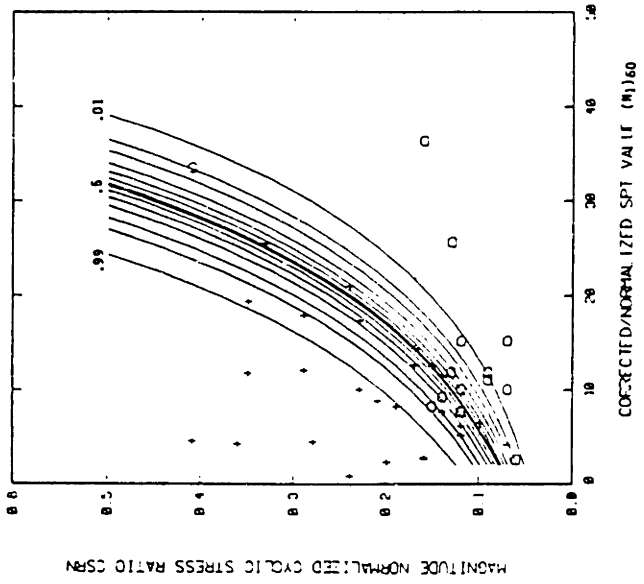
(c) Cases not from Japan/
California/China

Fig. 4.29 Contours of equal probability of liquefaction for clean sand based on Seed-Idriss model incorporating various indicator variables (fitted to all 278 cases).

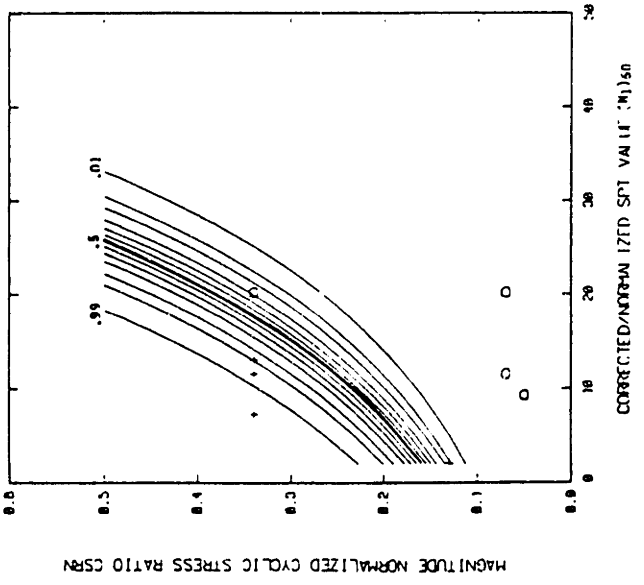
(Fig. 4.29 continued on next page)



(f) SPT performed using rope and drum

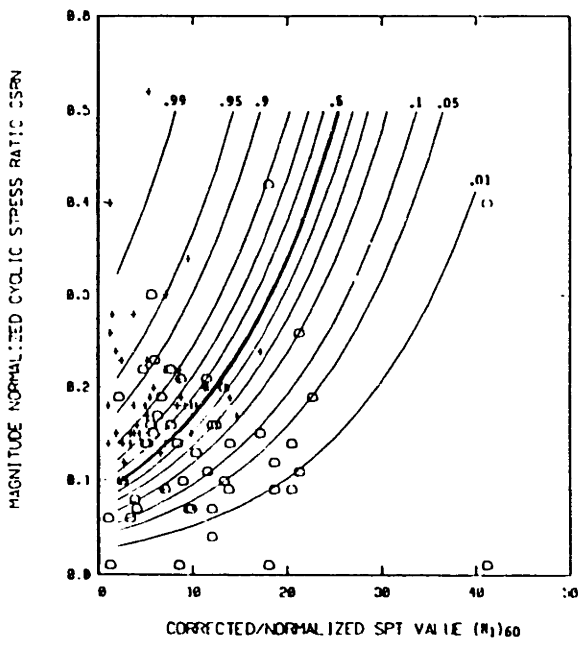


(e) Acceleration from intensity data

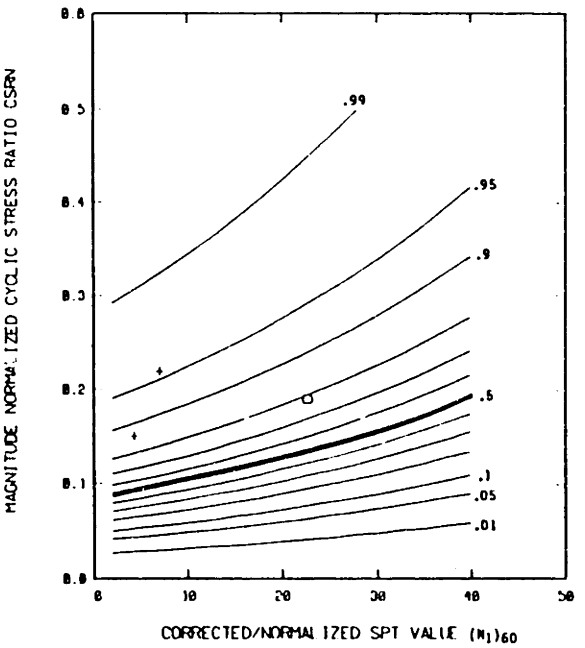


(d) Acceleration from site response analysis

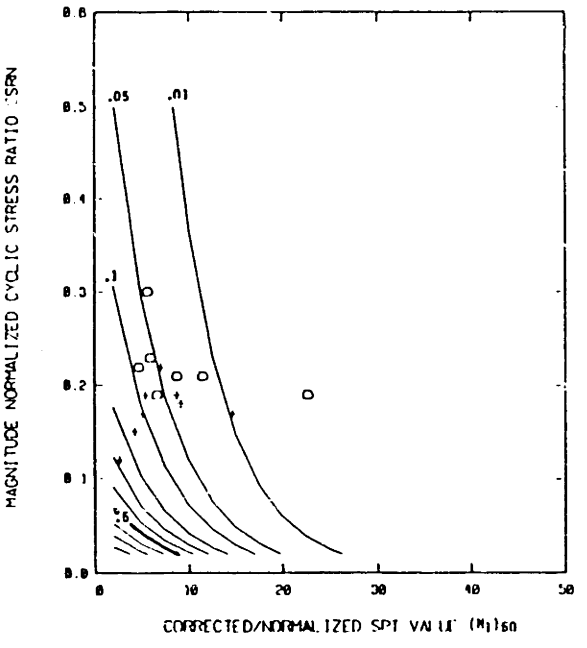
Fig. 4.29 (continued from previous page)



(a) Clean sand "base" model



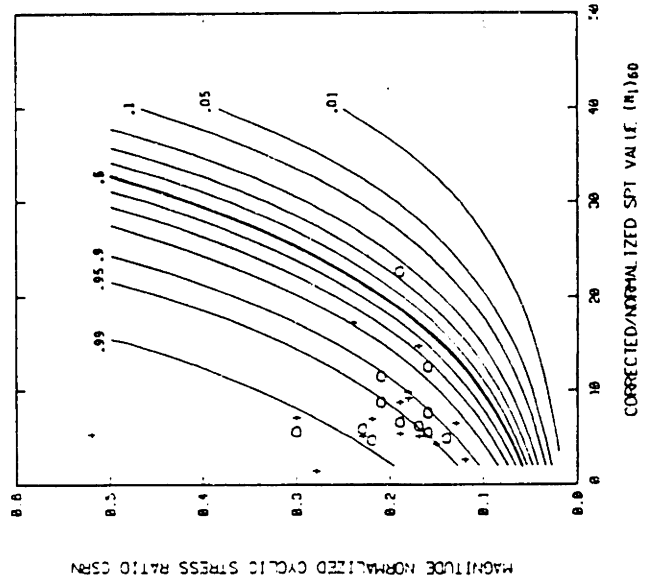
(b) Gravel Content
GC ≤ 10%



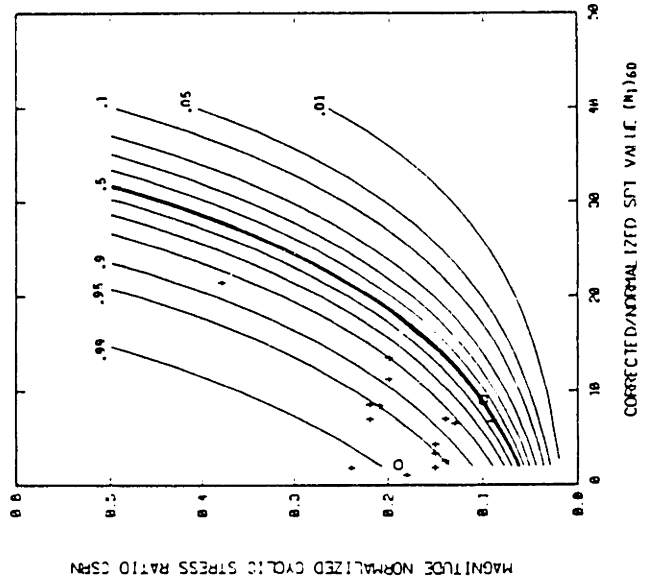
(c) Cases not from Japan/
California/China

Fig. 4.30 Contours of equal probability of liquefaction for silty sand based on Seed-Idriss model incorporating various indicator variables (fitted to all 278 cases).

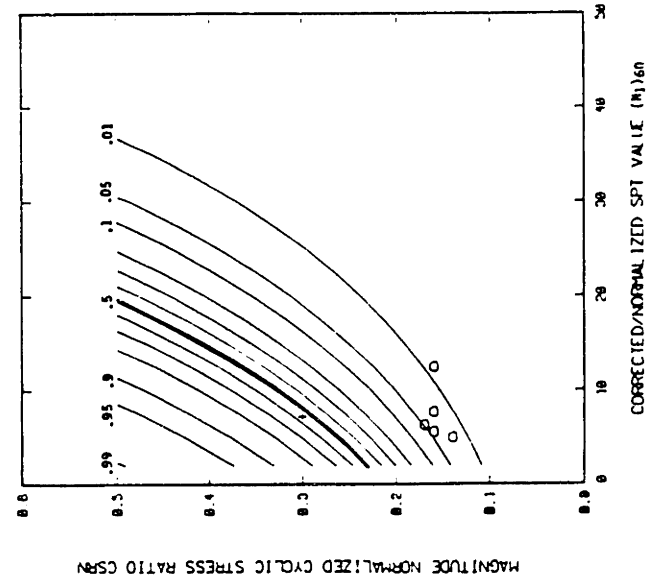
(Fig. 4.30 continued on next page)



(f) SPT performed using rope and drum



(e) Acceleration from intensity data



(d) Acceleration from site response analysis

Fig. 4.30 (continued from previous page)

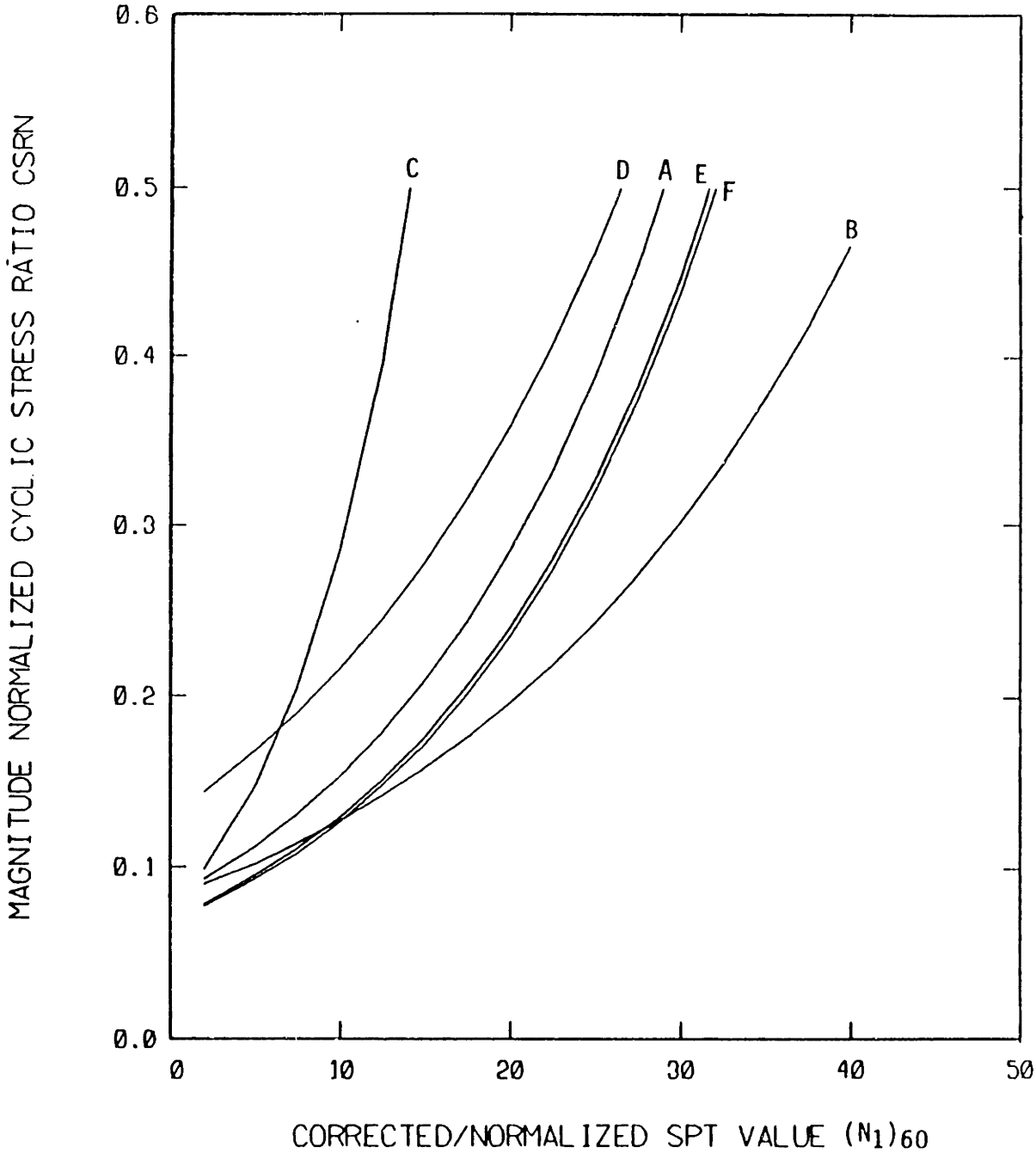


Fig. 4.31 Comparison of P = 0.5 contours for clean sand based on Seed-Idriss model incorporating various indicator variables (fitted to all 278 cases).

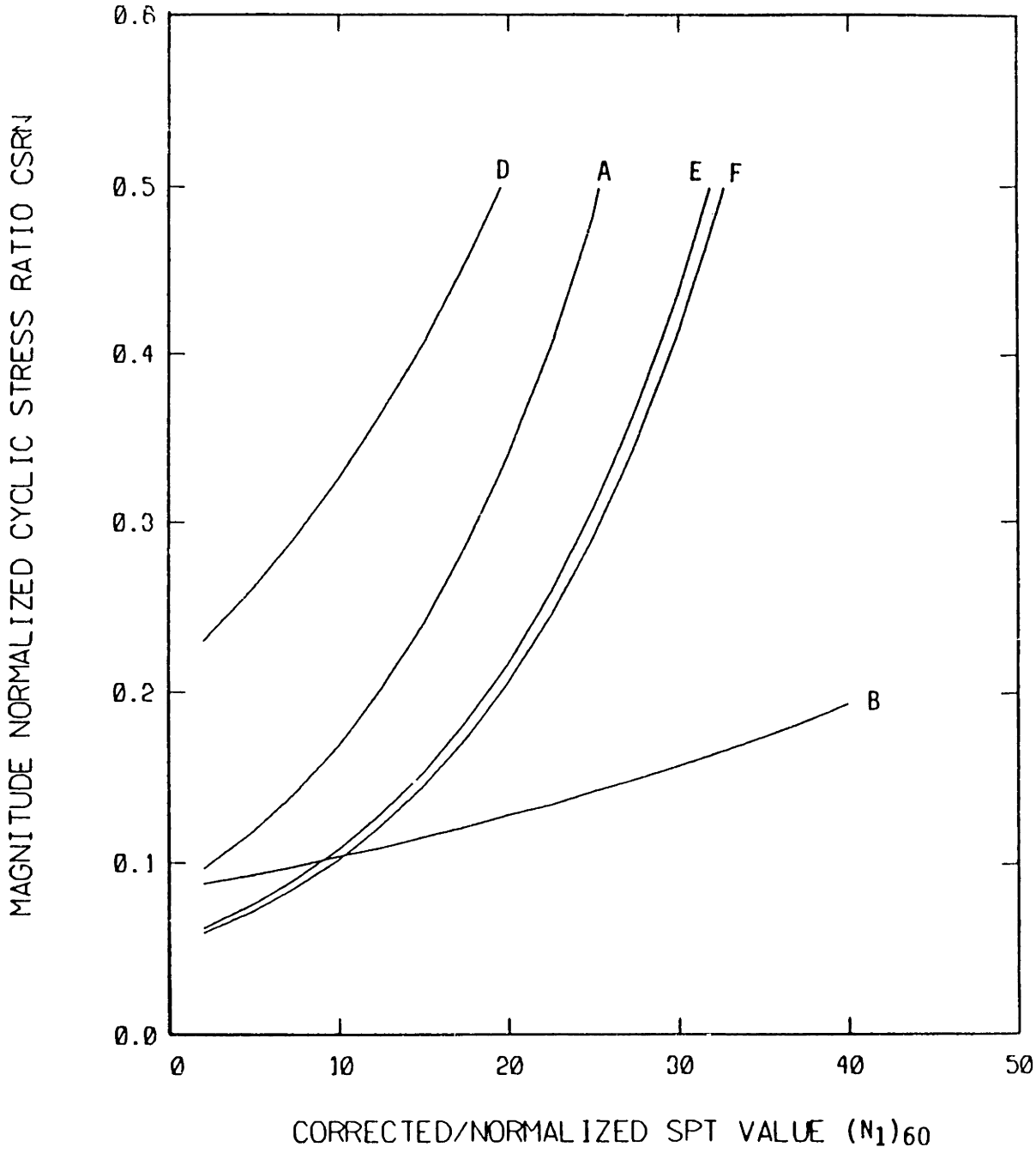
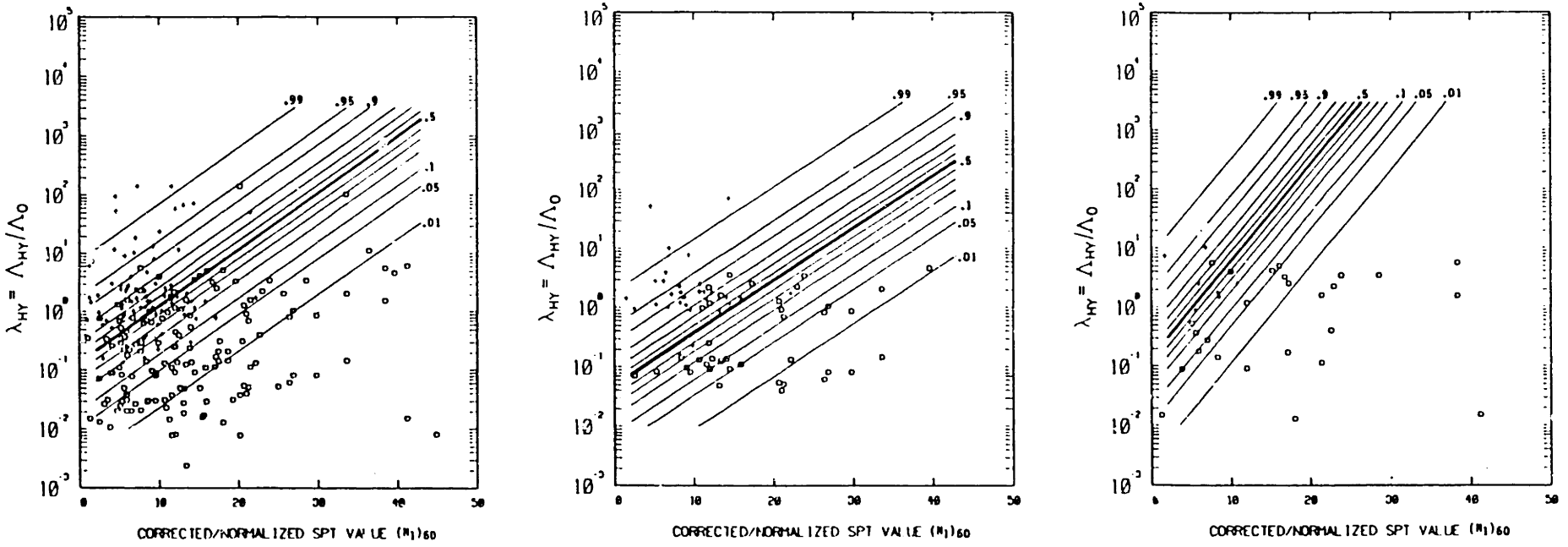


Fig. 4.32 Comparison of P = 0.5 contours for silty sand based on Seed-Idriss model incorporating various indicator variables (fitted to all 278 cases).



(a) "Base" model

(b) Method of obtaining representative N-value unknown

(c) Boring performed prior to earthquake

Fig. 4.33 Contours of equal probability of liquefaction for modified (hypocentral) Davis-Berrill model incorporating two indicator variables (fitted to 228 cases with focal depth data).

CHAPTER 5

NONPARAMETRIC BINARY REGRESSION: METHOD AND RESULTS

5.1 Introduction

In parametric methods of regression, it is generally assumed that the mathematical form of the relationship between the response and the explanatory variables are known. For example, in linear least squares regression with one explanatory variable x and response variable Y, the assumption of the mathematical form is a straight line:

$$Y = ax + b \tag{5.1}$$

where a and b are constants to be determined from analysis. Similarly, in binary logistic regression, we assume an equation of the form

$$\text{logit}(P) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \cdots \beta_m x_m \tag{5.2a}$$

or in vector notation:

$$\text{logit}(P) = \underline{x}^T \underline{\beta} \tag{5.2b}$$

where the vector of the coefficients $\underline{\beta}$ is obtained by fitting the equation to the data.

The objective of nonparametric methods of regression is to relax the constraint of having to assume a mathematical form as required in parametric analyses. Also, it may not always be clear whether the assumed form is correct or incorrect. Nonparametric methods allow us to

consider the question: Does the assumed mathematical form adequately model the data?

Although nonparametric estimators are more "flexible" and general, there is a trade-off associated with the increase in flexibility. In return, we lose the simplicity of having an explicit expression relating the response to the explanatory variables. The use of a simple parametric formulation is sometimes preferred especially when the regression is not the final end product, but simply a step along the way in the modelling of a more complex situation, e.g. seismic risk analysis. Thus the methods presented in this Chapter should be viewed basically as tools for exploratory data analysis.

The method of nonparametric binary regression used in this study is a "kernel" approach, which is formally described in Section 5.2. The basic idea behind the method can be illustrated by considering the data plot in Fig. 5.1, which typifies the Seed-Idriss method of liquefaction analysis. The crosses (+) indicate liquefaction (Y=1) and open circles indicate non-liquefaction (Y=0). There are regions in this plot that contain only crosses (+) or only circles (o), but there are also regions where the two mix. To estimate the probability of liquefaction (Y=1) at point A with coordinates $\underline{x}_A = (u_A, v_A)$, we consider a circular region or neighborhood (Circle A) centered at $\underline{x}_A$, and count the number of crosses (+) and the number of circles (o) within the region. The probability of liquefaction is estimated by the relative frequency:

$$\hat{P}(Y=1 \,|\, \underline{x}_A) = \frac{[\text{No. of (+)}] \text{ in Circle A}}{[\text{No. of (+)}] + [\text{No. of (o)}] \text{ in Circle A}} \qquad (5.3)$$

Similarly, the conditional probability of liquefaction can be evaluated using the same procedure at points B and C in Fig. 5.1. One would find $\hat{P}$ is about 0.5, 1.0 and 0.0 at points A, B, and C, respectively.

Instead of just counting the numbers of (+)'s and (o)'s within a region, an improvement to the estimate $\hat{P}$ would be to also assign weights to each data point, and have the weight decrease as the distance to the point of interest increases. For point A, those data closer to $\underline{x}_A$ would be assigned a larger weight than those farther away. The function that generates these weights is referred to as the "kernel" function.

Though the basic idea of kernel estimation is simple, there are several issues that must be considered for such an estimator of P. These include:

- What kind of smoothing or weight function should be used?
- How large a region ("bandwidth") around a data point should be considered?
- Should the region be circular or elliptical?
- If the region is elliptical, how should the principal axes be aligned?

The primary emphasis of this chapter will be on methdology and attempts at addressing the above issues. The data from the liquefaction case study catalog is used as a vehicle for developing the kernel estimators. Conclusions related to liquefaction risk analyses are discussed within the context of methdological developments. However, a summary of the practical results and implications is presented in Section 5.4.

5.2 Nonparametric Kernel Estimators

5.2.1 Basic Formulation

Let Y be a binary (0/1) response variable and $\underline{x}$ be a vector of m explanatory variables. Denote the set of given observations of $(Y,\underline{x})$ as $(Y_i,\underline{x}_i)$, i = 1, ... n, where n is the total number of observations. Using Eqn. (5.3), we introduced the concept of estimating the conditional probability $P(Y=1|\underline{x})$ by the ratio of counts of Y=1 to the total number of counts in a neighborhood $\mathcal{N}$ centered about $\underline{x}$. To formalize this concept, we rewrite Eqn. (5.3) as:

$$\hat{P}(Y=1|\underline{x}) = \frac{\sum_{i=1}^{n} \delta_i Y_i}{\sum_{i=1}^{n} \delta_i} \tag{5.4}$$

where δ_i is defined as:

$$\delta_i = \begin{cases} 1 \text{ if } \underline{x}_i \in \mathcal{N} \\ 0 \text{ otherwise} \end{cases} \tag{5.4b}$$

A more general class of estimators is obtained by replacing the indicator function δ_i by a smooth kernel function $K(\xi)$ which decreases gradually as ξ increases, where $\xi \equiv \xi(\underline{x},\underline{x}^*)$ is the distance measure between two points $\underline{x}$ and $\underline{x}^*$. Dencte $\xi_i = \xi(x,x_i)$ as the distance between $\underline{x}$ and the ith data point location $\underline{x}_i$. Then the probability of Y=1 (liquefaction) at $\underline{x}$ is estimated as:

$$\hat{P}(Y=1|\underline{x}) = \frac{\sum_{i=1}^{n} K(\xi_i)\, Y_i}{\sum_{i=1}^{n} K(\xi_i)} \tag{5.5}$$

This expression is exactly the form proposed by Nadaraya (1964) and Watson (1964), except that these authors consider Y to be a continuous rather than as a binary variable. Lauder (1983) also used an estimator of the type in Eqn. (5.5) for analyzing binary data.

Eqn. 5.5 can be further simplified to:

$$\hat{P}(Y=1|\underline{x}) = \sum_{i=1}^{n} w_i(\underline{x}) Y_i \qquad (5.6a)$$

where

$$w_i(\underline{x}) = \frac{K(\xi_i)}{\sum_{j=1}^{n} K(\xi_j)} \qquad (5.6b)$$

In this form, it is clear that $\hat{P}$ is simply a weighted linear combination of the observed responses Y_i, with weights that depend on the location $\underline{x}$ where $\hat{P}$ is calculated. An obvious property satisfied by the weights w_i is

$$\sum_{i=1}^{n} w_i(\underline{x}) = 1 \qquad \text{for any given } \underline{x} \qquad (5.7)$$

5.2.2 Kernel Function

There are relatively few restrictions on the choice of the kernel function. It is usually desirable for the kernel to be "smooth" and bounded, and $K(\xi)$ should be unimodal with its maximum value at $\xi = 0$. In applications involving kernel estimation of probability densities, $K(\xi)$ must also be normalized so that

$$\int_{-\infty}^{+\infty} K(\xi)d\xi = 1 \tag{5.8}$$

However, this property is not necessary in kernel regression, since any normalization constants applied to $K(\xi)$ cancel out in the ratio expressed in Eqn. 5.5. In general, any unimodal probability density function is acceptable.

In this study, the kernels were chosen to be proportional to a non-normalized multivariate normal probability density function of the form:

$$K(\xi) = \exp(-\frac{1}{2}\xi^2) \tag{5.9a}$$

where

$$\xi^2 = (\underline{x} - \underline{x}^*)^T \underline{S}^{-1} (\underline{x} - \underline{x}^*) \tag{5.9b}$$

and $\underline{S}$ is the analog of the covariance matrix, and will be referred to as the "smoothing matrix". The distance measure ξ is a statistical distance between the two points $\underline{x}$ and $\underline{x}$ rather than the more generally known metric of Euclidean distance. The measure ξ is also commonly referred to as the Mahalanobis distance.

It is generally accepted that the choice of the functional form of $K(\xi)$ is not critical for either regression or probability density estimation (see Rosenblatt, 1971; Hand, 1982; Härdle and Marron, 1985). The use of a Gaussian or normal kernel is fairly common, though it involves a slightly higher computational cost than other simpler kernels. However, the choice of the distance measure function ξ will significantly affect the results.

To clarify the application of the kernel estimator in the case two explanatory variables, we write the vector $\underline{x}$ as:

$$\underline{x} = \begin{bmatrix} u \\ v \end{bmatrix} \tag{5.10}$$

and the smoothing matrix $\underline{s}$ as:

$$\underline{S} = \begin{bmatrix} h_u^2 & \rho h_u h_v \\ \rho h_u h_v & h_v^2 \end{bmatrix} \equiv \begin{bmatrix} h^2 & \rho\gamma h^2 \\ \rho\gamma h^2 & \gamma^2 h^2 \end{bmatrix} \tag{5.11}$$

where h_u and h_v are the analogs of the standard deviations in the u and v directions and ρ is the analog of the correlation coefficient. [It is important to realize that h_u, h_v and ρ are only analogs of the standard deviation and correlation coefficient, and that they bear no relationship with the actual standard deviations and correlation coefficient of the data (u_i, v_i), $i = 1, \ldots n$.]

The quantities h_u and h_v in Eqn. 5.1 are commonly referred to as the "bandwidths" or "window widths" of the kernel. For convenience, we define $h_u = h$ and $h_v = \gamma h$ where γ is the ratio of the bandwidth in the v-direction to that in the u-direction. We will refer to ρ as the "orientation coefficient", since it basically controls the orientation of the kernel smoother. The squared distance between a generic location (u^*,v^*) to the location of the estimate (u°,v°) can now be written as:

$$\xi^2 = \begin{Bmatrix} u^* - u^\circ \\ v^* - v^\circ \end{Bmatrix}^T \begin{bmatrix} h^2 & \rho\gamma h^2 \\ \rho\gamma h^2 & \gamma^2 h^2 \end{bmatrix}^{-1} \begin{Bmatrix} u^* - u^\circ \\ v^* - v^\circ \end{Bmatrix} \tag{5.11a}$$

or

$$\xi^2 = \frac{1}{1-\rho^2}\left[\left(\frac{u^* - u^\circ}{h}\right)^2 + \left(\frac{v^* - v^\circ}{\gamma h}\right)^2 - 2\rho\left(\frac{u^* - u^\circ}{h}\right)\left(\frac{v^* - v^\circ}{\gamma h}\right)\right] \tag{5.11b}$$

The locus of points (u^*,v^*) for which ξ is constant defines an ellipse centered about (u°,v°) in the u-v plane as shown schematically in Fig. 5.2. All data points $\underline{x}_i = (u_i,v_i)$ which line on the ellipse are assigned equal weights w_i for the kernel estimate $\hat{P}(Y=1|\underline{x}=x^*)$ (Eqn. 5.6).

The parameter ρ affects the orientation of the principal axes of the ellipse as shown in Fig. 5.2. It also affects the area contained within the ellipse. If $\rho = 0$, the principal axes of the ellipse are aligned with the u and v axes and the area contained in the ellipse is at a maximum for fixed h and γ. In the limit as $\rho \to \pm 1$, the orientation of the ellipse will depend on h and γ, but the area contained in the ellipse would approach zero, as indicated in Fig. 5.2.

In the case where $\rho = 0$, the two-dimensional kernel $K(\xi)$ can be rewritten as the product of two univariate kernels:

$$K(\xi) = \exp\left\{-\frac{1}{2}\left(\frac{u-u^\circ}{h}\right)^2\right\} \exp\left\{-\frac{1}{2}\left(\frac{v-v^\circ}{\gamma h}\right)^2\right\} \tag{5.13}$$

The m-dimensional case, where the smoothing matrix has only non-zero diagonal elements, has the corresponding form:

$$K(\xi) = \prod_{j=1}^{m} K_j\left(\frac{\Delta_j}{h_j}\right) \tag{5.14}$$

where the index j refers to the j^{th} coordinate axis, Δ_j is the distance measure and h_j is the kernel bandwidth along that axis. This kernel function was originally proposed by Epanechnikov (1969) and is commonly referred to as a "product kernel".

Though product kernels are less general, they have the advantage of easier implementation, and appear to be favored in applications of

multivariate density estimation (Hand, 1982). While the bandwidth (diagonal) terms of the smoothing matrix $\underline{S}$ are easy to interpret physically, the analog covariance (off-diagonal) terms are much less so, particularly in dimensions higher than three. Furthermore, for product kernels, the number of independent smoothing parameters that need to be specified is equal to the dimensionality m, while for a general smoothing matrix, the number is m(m+1)/2. In later sections of this chapter, the differences between using product kernels ($\rho = 0$) and the more general formulation ($\rho \neq 0$) is investigated for the case of two explanatory variables. A kernel function using a smoothing matrix $\underline{S}$ with $\rho \neq 0$ will be referred to as an "oriented" kernel.

In the following two sections, we consider an extension and a modification of the kernel function as just described. The extension involves the idea of a location-dependent variable-size kernel, and is based on the work by Breiman, Meisel, and Purcell (1977). The modification involves the use of an integrated kernel function, which may be useful under conditions of monotonicity of the probability P, e.g. the liquefaction probability is known to increase with increasing values of the "load" variable (u=CSRN) and with decreasing values of "resistance" variable ($v=(N_1)_{60}$).

5.2.3 Variable Bandwidth Kernel

Consider again the two-dimensional kernel defined by the smoothing matrix $\underline{S}$ in Eqn. 5.11. One of the more critical parameters that defines the kernel is the bandwidth h. In classical kernel estimation, h is a fixed parameter independent of the location in the u-v plane. Keeping h

constant, however, may not be optimal since the density of points varies from location to location as shown, for example, in Fig. 5.1. A small bandwidth may be suitable in a region with a high density of data points, whereas a large bandwidth would be more appropriate in a region with sparse data.

Following Breiman et al. (1977), one may redefine h as:

$$h(\underline{x}) = c \cdot [d_k(\underline{x})]^b \tag{5.15}$$

where $d_k(\underline{x})$ denotes the distance from $\underline{x}$ to the kth nearest data point or "neighbor" $\underline{x}_i$ (i = 1, ... n). To be consistent in measuring near-neighbor distances, a Mahalanobis metric is used, i.e.:

$$d_k(\underline{x}) = [\underline{x}(k) - \underline{x}]^T \ \underline{S}_o^{-1} \ [\underline{x}(k) - \underline{x}]^T \tag{5.16}$$

where $\underline{x}(k)$ is the kth nearest data point to $\underline{x}$. A discussion of the procedures for establishing the analog covariance matrix $\underline{S}_o$ (different from $\underline{S}$ in Eqn. 5.9) is deferred to Section 5.2.7 on implementation. The parameters c and b in Eqn. 5.15 control the spread of the kernel; in particular, for b = 0, $h(\underline{x})$ degenerates to a fixed bandwidth.

In the paper by Breiman et al. (1977), a power coefficient b = 1.0 is used for probability density estimation. However, based on the work of Abramson (1982), Breiman (1985) has indicated that a better choice for parameter b is m/2, where m is the dimensionality of the space. Using a different approach, Silverman (1984) indicates that b = m/4 might be more appropriate. In this study, we have found that values of b between 0.6 to 0.8 ($b \approx m/3$) gave good results. Since both Abramson (1982) and Silverman

(1984) derive their conclusions from the asymptotic case when $n \to \infty$, whereas the sample used in this study is relatively small (182 data), the above discrepancy is not significant. Also, results using b = 0.5 or b = 1.0 are similar to those for b = 0.6 or b = 0.8.

5.2.4 Integrated Kernel

The purpose of the integrated kernel is to make better use of the data when there is prior information regarding the monotonicity of the function to be estimated. Consider the case of liquefaction analysis shown in Fig. 5.3. The quantity v = CSRN plotted on the vertical axis is an indicator of the intensity of earthquake loading. The quantity $u = (N_1)_{60}$ on the horizontal axis is an indicator of resistance to failure by liquefaction (Y=1). From physical considerations, we know that for any fixed value of u, increasing the load v causes an increase in the probability of failure. Conversely, for fixed v, increasing the resistance u decreases the probability of failure. These marginal directions of increasing P(Y=1) are shown in Fig. 5.3. Use of the kernels described previously do not guarantee monotonicity in the estimate of P. To do so, a modification is proposed such that a function analogous to a cumulative distribution is substituted for the normal density form of the kernel. A physical interpretation of the modification is presented below.

Consider a geographic site represented by a data point (+) where liquefaction (Y=1) was observed. Given the prior regarding the marginal directions of the monotonicity of P(Y=1), we can conclude that if the site were subjected to a higher intensity earthquake or had a lower resistance, an observation of Y=1 would also have occurred. The region

D_1 in Fig. 5.3 represents the combinations of u and v that would also have led to failure as implied by the single observed data point (+). Similarly, we can argue that the region D_0 determined by the point (o) in Fig. 5.3 represents more information on conditions u and v that would lead to non-liquefaction (Y=0) than would be conveyed by a local interpretation of the observation at point (o).

To incorporate the above idea, we propose a kernel estimator of the form:

$$\hat{P}(Y=1 \,|\, \underline{x}) = \sum_{i=1}^{n} w_i(\underline{x},\underline{Y}) Y_i \tag{5.17a}$$

where

$$w_i(\underline{x},\underline{Y}) = \frac{\int_{(D_1)_i} K(\xi)d\xi}{\sum_{i=1}^{n} \{Y_i \int_{(D_1)_i} K(\xi)d\xi + (1-Y_i) \int_{(D_0)_i} K(\xi)d\xi\}} \tag{5.17b}$$

Note that integration of the kernel function extends over the regions D_1 and D_0 as shown schematically in Fig. 5.3. The subscript i indicates the regions for each data point i=1,...,n. The interpretation of the integrations is that each observation of Y=1 or Y=0 represents a continuous multiplicity of additional observations over the respective regions D_0 or D_1. The integrals are an implicit summation of the contributions to the kernel weight $w_i(\underline{x},\underline{Y})$ of those multiplicities of implied data points.

A consequence of employing the integrated kernel is that the resulting estimate $\hat{P}$ is monotonic. This can be shown by examining the univariate analog representing a marginal estimate of P along the v-axis as shown schematically in Fig. 5.4. The contribution to weight functions $w_i(v,\underline{Y})$ of an arbitrary point (+) with response Y=1 is represented by the area

under the kernel function to the right of the point (+). Similarly, the contribution of a data point (o) with response Y=0 is the area under the kernel to the left of the data point. As the location of the estimate moves from the left to the right, the contribution of the Y=1 data point increases and the contribution of the Y=0 data point decreases. Thus each weight function $w_i(v,\underline{Y})$ increases monotonically and approaches $1/n_1$ in the limit as $v \to \infty$, where n_1 is the number of observations of $Y_i=1$. Similarly, $w_i(v,\underline{Y}) \to 0$ as $v \to -\infty$. Considering this in the context of Eqn. 5.17a, the marginal estimate $\hat{P}$ is evidently monotonic. Monotonicity in each of the marginal directions insures general monotonicity in the u-v plane.

An undesirable property of the integrated kernel estimator is that it is asymptotically biased. Consider again the case of one explanatory variable as shown in Fig. 5.4 and let the bandwidth of the kernel function approach zero so that the normal density kernel becomes effectively a Dirac delta function. In this case the integrated kernel estimate becomes in the limit as the number of data $n \to \infty$,

$$\hat{P}(Y=1|v) = \frac{n_{1L}}{n} \tag{5.18}$$

where n_{1L} is the count of occurrences of Y=1 to the left of v. This is, of course, different from the true probability P. The bias $(\hat{P} - P)$ is positive for large P and negative for small P, meaning that in the limit as $h \to 0$ and $n \to \infty$, $\hat{P}$ is a steeper function than P.

Although the integrated kernel estimator is asymptotically biased, we have found in the finite sample case that the bias is not severe, particularly when the size of the kernel is adjusted by an optimization

procedure that penalizes for the lack of fit of a biased estimate. In fact, the performance of the integrated kernel estimator in terms of bias and variance was found to be superior (for the liquefaction data) to that of the usual kernel estimator (see Section 5.3.5).

5.2.5 Moments of the Kernel Estimator

Consider the response Y_i observed at $\underline{x}_i$ as the outcome of a Bernoulli trial with probability masses $P(Y_i=1|\underline{x}_i) = P_i$ and $P(Y_i=0|\underline{x}_i) = 1 - P_i$. The mean and variance of Y_i are:

$$E[Y_i] = P_i \tag{5.19}$$

and

$$Var[Y_i] = P_i(1 - P_i) \tag{5.20}$$

For a non-integrated kernel, Eqn. 5.6 (repeated below) gives the estimator $\hat{P}$ as a linear combination of the variables Y_i:

$$\hat{P}(Y=1|\underline{x}) = \sum_{i=1}^{n} w_i(\underline{x})Y_i \tag{5.6}$$

Applying the properties of the expectation operator and using Equations 5.19 and 5.20, we obtain the results:

$$E[\hat{P}(Y=1|\underline{x})] = \sum_{i=1}^{n} w_i P_i \tag{5.21}$$

and

$$Var[\hat{P}(Y=1|\underline{x})] = \sum_{i=1}^{n} w_i^2 P_i(1-P_i) + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} w_i w_j Cov[Y_i,Y_j] \tag{5.22}$$

If the Bernoulli trails are independent, then Eqn. 5.22 simplifies to:

$$Var[\hat{P}(Y=1|\underline{x})] = \sum_{i=1}^{n} w_i^2 P_i (1 - P_i) \tag{5.23}$$

Furthermore, an estimate of the variance can be obtained by replacing P_i in Eqn. 5.23 with $\hat{P}_i$, i.e.:

$$\hat{Var}[\hat{P}(Y=1|\underline{x})] = \sum_{i=1}^{n} w_i^2 \hat{P}_i(1-\hat{P}_i) \tag{5.24}$$

where the values $\hat{P}_i$ are calculated using Eqn. 5.6, with the notation

$$\hat{P}_i \equiv \hat{P}(Y=1|\underline{x}_i) = \sum_{j=1}^{n} w_j(\underline{x}_i) Y_j \tag{5.25}$$

Eqn. 5.21 will be used in in Section 5.3.5 to study the bias of the estimator $\hat{P}$. Eqn. 5.23 provides a useful assessment of the variance of the estimator of P as a function of $\underline{x}$. In the case of the liquefaction data, observations Y_i may not be completely independent. As described in Chapter 3, the data collection procedure lends itself to a structure of uncorrelated clusters of data, but with some (positive) correlation among the observations within each cluster. For example, several observations are usually associated with the same earthquake and these observations may be correlated. Thus Eqn. 5.24 for is strictly not valid and provides only a lower bound estimate for the variance of $\hat{P}$. However, for the purposes of this Chapter, it is assumed that the intracluster correlations do not significantly affect the results. That this assumption is justified, is supported by the results of weighted logistic regression analyses described in Chapter 6.

Equations 5.21 and 5.24 are not valid for the integrated kernel proposed in Section 5.2.4. The reason is that the weights w_i for the integrated kernel are not independent of the responses Y_i. The variance and bias of the integrated kernel estimator were studied through Monte

Carlo simulation.

5.2.6 Optimization Criteria

Two criteria for the optimal selection of kernel parameters have been considered in this study, one based on maximization of a likelihood, and the other involving a least squares approach. In both cases, use is made of the principle of cross-validation. The first application of this idea in the context of kernel estimation was reported by Habbema, Hermans and Van Den Broek (1974), and it has since been used extensively (see Aitchison and Aitken, 1976; Aitken and MacDonald, 1979; Lauder, 1983).

There are methods for selecting kernel smoothing parameters other than cross-validation. The simplest relies on human judgment to determine the kernel parameters that adequately balance excessive smoothing versus excessive (high-frequency) "noise" in the estimate (Scott et al., 1976; Tapia and Thompson, 1978) or in the first and second derivatives of the estimate (Silverman, 1978). A more refined method involves the use of a maximum penalized likelihood, where the quantity to be maximized is defined as:

$$G = L - J$$

where L is the log-likelihood function and J is a penalty function that increases as the "roughness" of the estimate increases. This formulation was first proposed by Good and Gaskins (1971) and has been further developed by Tapia and Thompson (1978). However, the selection of an appropriate roughness penalty J may not always be straightforward. There are also several theoretical studies for the problem of probability density estimation that offer guidance on the selection of the smoothing matrix (particularly of the bandwidth h) as a function of the sample size

n; see Tapia and Thompson (1978) or Hand (1982) for a comprehensive review.

The log-likelihood L is written as:

$$L = \sum_{i=1}^{n} Y_i \ln \hat{P}_i + (1 - Y_i) \ln (1 - \hat{P}_i) \tag{5.26}$$

with $\hat{P}_i$ as given in Eqn. 5.25. The average squared error ASE is defined as:

$$ASE = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{P}_i)^2 \tag{5.27}$$

which is simply the mean value of the residuals of estimation, also referred to as the mean squared error. [The notation MSE is specifically avoided to prevent confusion between the quantity in 5.27 and the theoretical expected value of the squared error of the estimator, also commonly denoted as MSE in kernel estimation.] Both L and ASE are functions of the kernel smoothing parameters h, γ, ρ, etc. In parametric regression, one would normally proceed by either maximizing L or ASE with respect to the parameters to obtain an optimum solution. However, in the case of nonparametric kernel estimation, optimization would lead to a kernel in the form of a Dirac delta function, and hence to estimates $\hat{P}(Y=1 \,|\underline{x}) = 1$ for any $\underline{x} = \underline{x}_i$ when $Y_i = 1$, $\hat{P}(Y=1 \,|\underline{x}) = 0$ for any $\underline{x} = \underline{x}_i$ when $Y_i = 0$, and $\hat{P}$ undefined elsewhere. Cross-validation is an attractive alternative that circumvents this problem.

The cross-validation estimate of P(Y=1) at a data point location $\underline{x}_i$ (i=1, ..., n) is defined as:

$$P_i^* \equiv P^*(Y=1 \mid \underline{x}_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} w_j(\underline{x}_j) \quad Y_j \tag{5.28}$$

The fundamental difference between this estimate and the estimates in Eqn. 5.25 or Eqn. 5.26 is that the ith data point is excluded when estimating P at $\underline{x}_i$. The cross-validation log-likelihood function CVL is therefore defined as:

$$CVL = \sum_{i=1}^{n} Y_i \ln P_i^* + (1 - Y_i) \ln (1 - P_i^*) \tag{5.29}$$

This function is also frequently referred to as the pseudo-likelihood, the jackknife likelihood, or the deleted likelihood. Similarly, we define the cross-validation average squared error CVE as:

$$CVE = \frac{1}{n} \sum_{i=1}^{n} (Y_i - P_i^*) \tag{5.30}$$

Optimum kernel smoothing parameters can thus be obtained by either maximizing CVL or minimizing CVE. Several papers discussing the asymptotic optimality and consistency of cross-validated estimators have been published by Wong (1983), Chow, Geman, and Wu (1983), Rice (1984), Marron (1985) and Härdle and Marron (1985).

To illustrate the use of the cross-validation criteria, consider the optimization of a uniform size product kernel ($\rho = 0$) applied to the liquefaction data. Fig. 5.5 shows CVL and CVE as functions of the bandwidth h for a specified (optimized) value of bandwidth ratio γ. Also shown in the figure is the log-likelihood function L and the average

squared error ASE. The likelihood CVL achieves a maximum at $h \simeq 2.5$ and CVE achieves a minimum at $h \simeq 2.6$. In contrast, the maximum value of L and the minimum value of ASE occur both for $h = 0$, i.e. for a Dirac delta kernel function.

Fig. 5.6 illustrates the effects of using non-optimal bandwidths in the calculation of $\hat{P}(Y=1|\underline{x})$. In particular, Fig. 5.6(a) shows an estimate where h is relatively small and smoothing is too local. In contrast, Fig. 5.6(c) shows an estimate where h is large and there is too much smoothing. The optimal solution for fixed-size product kernel is shown in Fig. 5.6(b). It should be noted that even this "optimum" kernel estimate has features that are questionable, including the non-monotonicity especially evident in the upper middle portion of Fig. 5.6(b). Also, the spread of contours of $\hat{P}$ in the lower left corner of the figure appear too large. These features are partly due to the sparseness of data and the uneven density of data points. Thus, variable size and integrated kernels were used to try to compensate for these problems.

Optimization of CVL or CVE was performed through a hierarchical procedure described in the following section, using the IMSL subroutine ZXMIN. It was empirically observed that, for the data analyzed, both the functions CVL and CVE are usually globally concave (either downwards or upwards) and that the optimization procedure converged without problems. Exceptions occurred under certain conditions for the general kernel smoother ($\rho \neq 0$) and also for optimization of the parameter k which determines the distance d_k for the variable kernel smoother. These exceptions are discussed within the context of the results in Section 5.3.

The difference between using the likelihood and the least-squares cross-validation approaches has been investigated by Rudemo (1982) and Marron (1985) within the context of density estimation. Based on the results of various Monte Carlo simulations, Rudemo (1982) observed that optimization of cross-validated likelihood produces estimators that are more sensitive to outliers or isolated data, compared with the least squares approach. Marron (1985) derives asymptotic results that show reasons for preferring the least squares formulation. For the liquefaction data used in this study, differences between the maximum CVL and the minimum CVE smoothing parameters are small. A typical comparison of the optimum CVL and CVE solutions is shown in Figures 5.7(a) and 5.7(b). Note that the only detectable difference is in the upper region, where the data points are relatively sparse. The good agreement is due to the absence of conspicuous outliers in the data.

In the context of binary regression, outliers are data points $(\underline{x}_i, Y_i)$ such that $Y_i = 1$ if $P(\underline{x}_i)$ is close to 0, or $Y_i = 0$ if $P(\underline{x}_i)$ is close to 1. In cases where the outliers are significant and a choice has to be made between the likelihood and squared error formulations, a basic philosophical question must be confronted: Are the outliers wrong data or are they valid observations? If one believes that they are simply wrong observations, then one should reduce their effects by using a least squared error approach. However, if one believes that the data is valid, a likelihood formulation may be more appropriate.

Although initially, both the maximum CVL and minimum CVE approaches to selecting kernel smoothing parameters were considered, the maximum CVL approach was finally preferred. The choice was based on the fact that

logistic regression models had been fitted in Chapter 4 using a maximum likelihood approach. Therefore, the CVL approach afforded a more consistent basis for comparing parametric and nonparametric results.

5.2.7 Notes on Implementation

The preceding sections have described a variety of 2-dimensional kernel estimators of $P(Y=1|\underline{x})$, ranging from uniform size product kernels with only two parameters to variable size oriented kernels with six parameters (three smoothing parameters plus three parameters that define the metric for near neighbor distances). When optimization is involved, the cost of computing increases considerably as more complex kernels are used. In terms of sample size, the cost is roughly proportional to n^2. A comparison of the relative costs (in terms of CPU time) of implementing various kernels is shown in Table 5.1. [Note that for future applications, more efficient optimization algorithms than ZXMIN could considerably reduce computation time.] Out of necessity, a hierarchical optimization procedure was developed that minimizes the total number of parameters that have to be determined at any one time.

Consider for example the integrated variable kernel with the following parameters: the neighbor index k, the constants c and b in Eqn. 5.15, the bandwidth size constants, and the covariance matrix $\underline{S}_o$ with 3 parameters, which is needed to calculate d_k using Eqn. 5.16. The hierarchical optimization procedure works as follows:

1. Determine the optimum bandwidth h and bandwidth γ for a uniform size product kernel ($\rho = 0$).
2. If an oriented kernel is desired, optimize γ and ρ considering h fixed, with h determined from Step 1. Otherwise, go to Step 3.

3. Use the parameters h, γ and ρ determined in Steps 1 and 2 to define $\underline{S}_o$ as:

$$\underline{S}_o = \begin{bmatrix} h^2 & \rho\gamma h^2 \\ \rho\gamma h^2 & \gamma^2 h^2 \end{bmatrix} \tag{5.31}$$

Determine the optimum k by trying several values of k, and optimize c and b for each k. [This is recommended because k is an integer.] At this step, one obtains the optimum parameters for the variable kernel.

4. Using the optimum value of k from Step 3, optimize the c and b for the integrated kernel. Note that it may be desirable to also fix b, choosing it to be the optimum obtained in Step 3 and only optimize c.

The above procedure was found to be fairly successful. Even though the absolutely optimal kernel is not obtained, the estimator $\hat{P}$ is probably fairly close to the optimum. Also, the accuracy required in the estimation of the parameters is not high because the function $\hat{P}(Y=1|\underline{x})$ is fairly insensitive to small changes in the parameters. Two digit accuracy was found to be sufficient for analysis of the liquefaction data. The reason for optimizing the parameters of the oriented kernel in two steps instead of one has to do with the characteristics of the CVL and CVE functions, which will be discussed in Section 5.3.4. It should be noted that the integrated kernel could be used in any of the steps above. Also, depending on the data set, it may not be necessary to implement the most sophisticated smoother; in particular for very large data sets, the uniform product kernel may give sufficiently good results.

To obtain graphical presentations of the $\hat{P}$ estimates, such as those shown in Figures 5.6 and 5.7, $\hat{P}$ was calculated on a grid of points as

shown in Fig. 5.8. Then linear interpolation between the grid points was used to draw the contour lines of $\hat{P}$.

5.3 Analysis of the Liquefaction Data

5.3.1 Preliminaries and Outline

The data set chosen for analysis is a 182 case subset of sites with clean sands (fines content FC < 12%) from the larger liquefaction catalog of 278 cases described in Chapter 3 and Appendix A. This data subset has already been analyzed in Chapter 4 using logistic regression. The data was found to be fitted very well by the following logistic model with two explanatory variables:

$$P(Y=1) = 1/\{1 + \exp -[\beta_o + \beta_1 \ln(CSRN) + \beta_2(N_1)_{60}]\} \qquad (5.32)$$

with estimated parameters $\hat{\beta}_o = 16.447$, $\hat{\beta}_1 = 6.4603$, and $\hat{\beta}_2 = -0.3976$. The associated log-likelihood for the fitted model is $L(\underline{\hat{\beta}}) = -44.705$ and the log-likelihood for $\beta_o = \beta_1 = \beta_2 = 0$ is $L(0) = -126.15$. The modified likelihood ratio index (a goodness-of-fit indicator) is MLRI = 0.646.

The question naturally arises as to whether Eqn. 5.32 and similar equations fitted in Chapter 4 are of the appropriate parametric form. Kernel estimation can be used to judge the adequacy of particular parametric assumptions. Specifically, we shall compare the logistic model (Eqn. 5.32) with the nonparametric kernel estimates described here.

This section presents results obtained using various kernel functions described in Section 5.2. The discussion will concentrate primarily on product kernels ($\rho = 0$). Oriented kernels ($\rho \neq 0$) will be considered separately in Section 5.3.4. Starting with the simple product kernel with

two parameters h and γ, we shall use increasingly more complex kernel functions, illustrating the features and/or improvements of each step of sophistication.

5.3.2 Choice of Smoothing Space

The choice of a "smoothing space" is a preliminary step of the analysis. For example, one might use the kernel estimator in either the space of CSRN and $(N_1)_{60}$ or of log(CSRN) and $\log[(N_1)_{60}]$ prior to smoothing. This problem is the analog of having to choose the parametric form of the explanatory variables in linear logistic regression (see Chapter 4, Section 4.4.3). [Note: The notational convention used here is that "ln" is used to denote the natural logarithm, whereas "log" refers to the base 10 logarithm.]

Fig. 5.9 shows the optimum estimates of P using uniform product (UP) kernels in the four smoothing spaces that result as possible combinations of logarithmically transforming one or both of the original variables. For the purpose of comparison, the results of Fig. 5.9 are shown in natural coordinates in Fig. 5.10. It is readily seen from this figure that the $\hat{P}$ estimates for the uniform size product (UP) kernel are fairly insensitive to the choice of smoothing space. Robust features which are apparent in each of the four estimates are:

1. A positive intercept of the vertical CSRN coordinate axis by the equi-probability lines.
2. A generally concave-upward curvature of the equi-probability lines.

The corresponding parametric results from logistic regression are shown in Fig. 5.11, and display much higher variability. This is to be

expected since the kernel method is a "local" procedure, whereas parametric estimates are implicitly "global".

Another comparison of the nonparametric results is shown in Table 5.2, which gives the optimum smoothing parameters h and γ and the associated log-likelihood statistics for the optimal estimates. The highest log-likelihood L corresponds to the choice of $u = (N_1)_{60}$ and v = log(CSRN). The highest logistic regression log-likelihood $L(\hat{\underline{\beta}})$ also corresponds to this choice of smoothing space. However, the smoothing space with highest value of CVL does not correspond to that with the highest value of L, though the differences between the highest and second highest value of CVL is negligible. Both the CVL and L statistics are given for the various kernel estimates presented in this Chapter, though the CVL statistic was preferred for comparing the different estimates. Informal comparisons were also made between the likelihood statistic L of the kernel estimates with $L(\hat{\underline{\beta}})$ from logistic regressions. However, in order to make formal comparisons, one needs to know the equivalent degrees of freedom of the kernel estimate.

The different natures of kernel and parametric estimates are also reflected in the likelihoods shown in Table 5.2. The difference between the highest and lowest $L(\hat{\underline{\beta}})$ (best and worst fits) is 10.77 for the parametric solutions. The same difference in L is only 0.92 and in CVL is 3.51 for the kernel estimates. Thus, despite an inappropriate choice of the smoothing space, the kernel estimate, because of its "local" nature, can accommodate the poor choice of variables better than a parametric model.

5.3.3 Product Kernel Estimates

Estimates of P(Y=1) using product kernel functions in the plane of u = $(N_1)_{60}$ and v = log(CSRN) are shown in Fig. 5.12. The smoothing parameters and the log-likelihood statistics for these estimates were obtained using the optimization procedure described in Section 5.2.7 and are given in Table 5.3. There is an improvement on the log-likelihood statistics (CVL and L) as one goes from the uniform (UP) kernel to the variable (VP) kernel to the integrated variable (IVP) kernel. The more significant improvement occurs for the integrated (IVP) kernel. As one can see in Fig. 5.12, the improvements due to the variable (VP) kernel occur primarily in the region of sparse data. Using the integrated variable (IVP) kernel produces further smoothing and narrowing of the $\hat{P}$ contours. Similar trends are seen in Fig. 5.13, which shows the estimates $\hat{P}$ obtained using a suboptimal choice of the smoothing space [u = $(N_1)_{60}$, v = CSRN].

Some observations can also be made about the estimates from variable kernels. First, the CVL function is not unimodal or globally concave with respect to the near neighbor indicator k, as shown in Fig. 5.14(a). This is attributed to the fact that k is an integer variable, and for the relatively small sample used here, the kth nearest neighbor distance d_k is somewhat erratic. However, the differences in the CVL function are not large and are visually exaggerated by the choice of scale in Fig. 5.14(a). Fortunately, the optimum values of b and c are relatively insensitive to the value of k (see Fig. 5.14(b) and 5.14(c)). The parameters b and c also tend to compensate for suboptimal choices of k. This is more clearly illustrated in Fig. 5.15 where the average distance to kth nearest

neighbor $\bar{d}_k$ (averaged over all data points) is plotted against k [in Fig. 5.15(a)]. While $\bar{d}_k$ increases monotonically with k, the quantity $c(\bar{d}_k)^b$ remains fairly constant, as shown in Fig. 5.15(b). A similar observation was made by Breiman et al. (1977). The optimum value of $c(\bar{d}_k)^b$ corresponds approximately to the optimal bandwidth h for a uniform size kernel. The optimum values of b are in the range of 0.6 to 0.8, tending towards 0.8 for larger values of k. These values are intermediate between those recommended by Abramson (1982) and Breiman (1985) and by Silverman (1984), based on asymptotic results. However, the estimates $\hat{P}$ are not very sensitive to the parameters k or b, as shown by a series of comparative results in Fig. 5.16 and Fig. 5.17.

Integrated uniform (IUP or IUV) kernels were also considered. The optimum parameters and the likelihood statistics for these kernels are listed in Table 5.4. Their performance is very similar to that of integrated variable kernels and their likelihood statistics are only slightly inferior. This suggests that, particularly in situations where large amounts of data are avaiable, the choice of using a variable or uniform integrated kernel may lead to practically equivalent estimates.

5.3.4 Oriented Kernel Estimates

Several problems have been encountered in the optimization of oriented kernel parameters. One is illustrated by Fig. 5.18(a) which shows the cross-validation likelihood CVL as a function of the orientation parameter ρ. The values of CVL, h and γ shown in Fig. 5.18 were obtained using a uniform size kernel and maximizing CVL for a fixed value of ρ. The CVL function exhibits a local maximum at $\rho \simeq 0.8$, but a global maximum at

$\rho \rightarrow 1.0$. The dashed lines in Fig. 5.18 do not represent discontinuous jumps, but rather the fact that the optimization algorithm converged to either the local or global maximum depending on the initial guess of the parameters. The apparent discontinuity is a result of plotting a multivariate function in two dimensions. Such problems were also observed in the CVE approach.

The global maximum CVL solution, in which the kernel smoothing parameters $\rho \rightarrow 1$ and the bandwidth $h \rightarrow \infty$, represents a degenerate function with straight parallel lines. It is proposed that the local maximum at $\rho \simeq 0.8$ indicates a resistance by the data to be characterized by such a function. Depending on the choice of smoothing space, a local maximum in the CVL function may not exist, as shown in Fig. 5.19. The difference between the kernels used in Fig. 5.18 and Fig. 5.19 is that the former smooths in the space of $u = (N_1)_{60}$ and v = CSRN and the latter in the space of $u = (N_1)_{60}$ and $v = \log(\text{CSRN})$. A common feature in both spaces is that the bandwidth h remains relatively constant over a large range of ρ and this constant value is approximately equal to the optimum h for $\rho = 0$. This observation has led to the step in the hierarchical optimization scheme of Section 5.2.7 where the optimum h and γ are first obtained assuming $\rho = 0$; then h is fixed and new optimum values of ρ and γ are found. In the case where a local optimum in the CVL function exists, this procedure produced results that are practically identical to those at the local optimum.

The argument for the reason why there is a local maximum in the CVL function for one smoothing space and not the other is shown by the comparison presented in Fig. 5.20. In part (a) of this figure the results

of smoothing in the space [u = $(N_1)_{60}$, v = CSRN] illustrate that the shapes of the equi-probability lines vary in slope with the characteristic concave upward shapes discussed previously in Section 5.3.2. However, in the smoothing space [u = $(N_1)_{60}$, v = log(CSRN)] shown in Fig. 5.20a, the equi-probability contours are approximately straight lines. Thus in this space, the interpretation of $\rho \rightarrow 1$ as optimum is that the best solution is obtained by smoothing along the straight lines of equal liquefaction probability. In the "natural" smoothing space represented by Fig. 5.20(a), since the equi-probability contours are not inherently straight, a compromise value of ρ occurs and results in the local maximum value of CVL. There is an improvement of the log-likelihood statistics for the oriented kernel (at this local optimum) over the product kernel (ρ = o) as shown in Table 5.3. Furthermore, this solution has a plausible physical interpretation, as described above, and thus is considered to be more resonable than the global maximum CVL solution in the natural space. Fig. 5.20 also illustrates a potential source of bias that is a feature of the oriented kernel: the lines of equal probability tend to align themselves along the direction of the major principal axis of the smoothing ellipse. This factor should be considered in the interpretation of oriented kernel estimates.

The use of variable oriented kernels and integrated oriented kernels produces improvements in some cases. For the oriented kernel in the [u = $(N_1)_{60}$, v = log(CSRN)] space, there was no increase in the CVL function in going from a fixed to a variable kernel; in fact, the optimum value of b was zero regardless of the choice of k. For the [u = $(N_1)_{60}$, v = CSRN] smoothing space, an improvement occurs in going from the uniform to the

variable size smoother, but the likelihood statistics are not as good for the integrated variable kernel as shown in Table 5.3. The integrated uniform oriented (IUO) kernels performed approximately the same as the integrated product (IVP) kernels as shown in Table 5.4. The differences between the uniform, variable, and integrated variable oriented kernels are shown in Fig. 5.21 for the "natural" smoothing space. Fig. 5.21 should be compared to Fig. 5.13 to view the differences in estimates obtained from product and oriented kernels.

For smoothing in the [$u = (N_1)_{60}$, $v = \log(CSRN)$] space, the integrated uniform oriented (IUO) kernel and the integrated variable product (IVP) kernel were practically identical in terms of likelihood statistics (see Tables 5.3 and 5.4). A comparison of these two results in Fig. 5.22 indicates the difference to be that the contour lines for $\hat{P}$ are closer together for the IUO kernel. This figure should also be comapred with Fig. 5.11(c) which shows the contour lines from the best-fit logistic regression estimate. The log-likelihood of the logistic fit is $L(\hat{\underline{\beta}}) = -44.71$, whereas the nonparametric log-likelihoods are $L = -44.14$ and $L = -44.13$ for the IVP and IUO kernels. The similarities between the kernel estimates and the logistic regression estimates are striking and suggest that the logistic regression model is appropriate for the data.

5.3.5 Standard Deviation and Bias of the Kernel Estimators

The standard deviation and the bias of the estimators of $P(Y=1|\underline{x})$ are denoted as $SD[\hat{P}]$ and $BIAS[\hat{P}]$. The bias of the estimator is defined as:

$$BIAS[\hat{P}] = E[\hat{P}] - P \tag{5.33}$$

where $E[\hat{P}]$ denotes the expected value of $\hat{P}$. Technically speaking, to obtain $E[\hat{P}]$, $SD[\hat{P}]$, and $BIAS[\hat{P}]$, the "true" underlying function $P(Y=1|\underline{x})$ must be known. In this section, we study the properties of kernel estimators under the assumption that $P(Y=1|\underline{x})$ is the logistic function given in Eqn. 5.32. Because the true function probably does not deviate much from that in Eqn. 5.32, the estimates of $E[\hat{P}]$, $Var[\hat{P}]$, and $BIAS[\hat{P}]$ should be good approximations to the true values.

The standard deviation and bias of the kernel estimators are a function of the set of data $\underline{x}_i = (u_i, v_i)$ $(i = 1,\ldots,n)$ and the corresponding probabilities P_i. In this study, we used the $\underline{x}_i$ corresponding to the liquefaction data analyzed in previous sections of this chapter. Calculation of $E[\hat{P}]$ and $SD[\hat{P}]$ for the non-integrated kernel estimators were performed using the formulas presented in Section 5.2.5. For the integrated kernel, estimates of $E[\hat{P}]$ and $SD[\hat{P}]$ were obtained through Monte Carlo simulation, the details of which are presented later.

The values of $E[\hat{P}]$, $SD[\hat{P}]$, and $BIAS[\hat{P}]$ for three kernel estimators are shown in Figures 5.23, 5.24, and 5.25. The kernels involved in the sequence of figures are the uniform product (UP), variable product (VP) and the integrated variable product (IVP) kernels, using the $[u = (N_1)_{60}, v = \log(CSRN)]$ smoothing space. The contour lines of $E[\hat{P}]$ shown in part (a) of these figures should be compared to the results of Fig. 5.12. Notice that the agreement between $E[\hat{P}]$ and $\hat{P}$ based on the data is fairly good. Part (b) of Figures 5.23, 5.24, and 5.25 show a marked decrease in the standard deviation of the estimates as we progress from the UP to the VP to the IVP kernels. This is especially true in the regions with sparse

data. Similarly, there is a decrease in the general level of bias as shown in part (c) of the figures. It is evident that the best estimator is obtained using the IVP kernel shown in Fig. 5.25. Note especially the low values of SD[$\hat{P}$] and the nearly zero values of BIAS[$\hat{P}$] in the region of relatively high density of data (in the lower left corner of Fig. 5.25(c)). High values of bias exist only in the region of relatively sparse data, where the accuracy of the estimate for most practical applications is not extremely important.

For all three kernel estimators, the maximum SD[$\hat{P}$] occurs near the P = 0.5 contour line, which is what we would expect for a sum of Bernoulli random variables. The contour line of zero bias is also near the P = 0.5 line, with generally negative bias for P > 0.5 and positive biases for P < 0.5. This can be seen more clearly in Fig. 5.26 which shows plots of BIAS[$\hat{P}$] versus P, calculated at the data points $\underline{x}_j$. There is a clear sinusoidal pattern to the data shown in Parts (a) and (b) of this figure, which correspond to the non-integrated UP and VP kernel estimators. The bias is nearly zero for P = 0, P = 0.5 and P = 1.0. This pattern is expected for standard kernel estimators and corresponds to the flatter slope of the estimate $\hat{P}$ relative to the true P. Fig. 5.26(c) does not show this sinusoidal pattern and further illustrates the dramatic reduction of bias that is obtained with integrated kernel smoothers. As mentioned in Section 5.2.4, integrated kernel estimator is asymptotically biased, with generally negative bias for small P (P < 0.5) and positive biases for large P (P > 0.5). This is the exact opposite of the bias for the non-integrated kernels for relatively small samples. Thus the nearly zero bias of the integrated kernel estimator may be thought of as the

counter-balancing of the two bias effects in opposite directions.

The results shown in Fig. 5.25 and Fig. 5.26(c) for the integrated kernel were obtained through Monte Carlo simulation. As noted in Section 5.2.5, Equations 5.21 and 5.23, which were used to obtain the standard deviation and bias results for non-integrated kernels, are not valid for integrated kernels. The number of simulations used for the integrated kernel results is 30. This number was selected following a study of similar simulation results for the uniform product (UP) kernel, which required considerably less (one order of magnitude) computer time. These results are shown in Figures 5.27 and 5.28 compared with the theoretical values from Equations 5.21 and 5.23. While the results for 60 simulations are slightly better than those for 30, the estimates from the 30 simulations are already very accurate.

Estimates of the standard deviaion of $\hat{P}$ were also obtained for the actual data using Eqn. 5.24, and these estimates are denoted as $\widehat{SD}[\hat{P}]$ (rather than $SD[\hat{P}]$). Results for the non-integrated UP and VP kernels are shown in Fig. 5.29, which should be compared with part (b) of Figures 5.23 and 5.24. The $\widehat{SD}[\hat{P}]$ values and shapes of the contour lines are very similar to those of $SD[\hat{P}]$ obtained assuming that P is the logistic function in Eqn. 5.32 and Figure 5.11(c).

As a further analysis, the expectations $E[\hat{P}]$ were calculated using the other logistic functions shown in Fig. 5.11 as the function P. The corresponding estimates for each of these assumed underlying functions are shown in Fig. 5.30, indicating that the expected shapes of the equi-probability contours would be significantly different in the various

cases. This result again re-emphasizes the conclusion that the proposed logistic function is a fairly good model for the data of the present liquefaction catalog.

5.4 Chapter Summary and Conclusions

In Chapter 4, various parametric fits to the liquefaction data were obtained using logistic regression. The four parameteric forms shown in Fig. 5.11(a) through (d) were considered as possible logistic models. However, based on physical considerations and goodness-of-fit statistics, it was concluded that the form of model represented in Fig. 5.11(c) is the most appropriate. Further confirmation of this conclusion has been found in the present chapter, through the use of nonparametric binary regression. Four features of the regression fit in Fig. 5.11(c) were of concern:

1) The positive intercept on the vertical CSRN axis by the equi-probability lines.
2) The concave-upwards curvature of the contour lines.
3) Symmetry of the contours about the $\hat{P} = 0.5$ line.
4) Divergence of the contour lines as CSRN and $(N_1)_{60}$ increase.

To explore questions regarding whether the above features are the result of the parametric logistic assumption or are actually embedded in the data, non-parametric estimation methods have been used. A family of kernel estimators were developed ranging from classical uniform product kernels to more sophisticated variable, integrated, and oriented kernels. Features (1), (2) and (3) above were confirmed using uniform (UP) and variable (VP) product kernels. From simulation studies and analytical

approximations, it was concluded that the UP and VP kernel estimators tended to be biased with negative biases for large P and positive biases for small P. The use of integrated and/or oriented kernels significantly reduces this bias and produced improved estimates for the data set considered. However, other types of bias can occur for the integrated and oriented kernels and these should also be considered in applying the kernel estimation procedures to other data sets.

Optimum kernel parameters were obtained by a cross-validation method. Problems with optimization of oriented kernel parameters were encountered and mitigated by a hierarchical procedure proposed in this chapter. In contrast, no optimization problems were encountered using product kernels. Also, the combination of features represented by the integrated product kernel produced results that are comparable to any using oriented kernels. For these reasons, the use of product kernels was favored in our analyses.

It would have been desirable to apply the proposed kernel estimation methods to other data sets or subsets of the liquefaction catalog as this was done in Chapter 4 with logistic regression. However, the computational costs of the kernel estimators are prohibitive compared to parametric logistic regression. Nevertheless, based on the present results, it is concluded that the proposed logistic regression model is appropriate and fairly accurate in representing the features of the liquefaction data. For the clean sand data, the logistic model is written as:

$$P = 1/[1 + \exp - (\underline{x}^T\underline{\beta})] \tag{5.34a}$$

where

$$\underline{x}^T\underline{\beta} = 16.447 + 6.4603\ \ln(CSRN) - 0.3976\ (N_1)_{60} \tag{5.34b}$$

The best nonparametric results are considered to be those obtained using the IVP and IUO kernels as shown in Fig. 5.22. The $\hat{P}$ = 0.1, 0.5, and 0.9 lines of equal probability for these three estimates are compared with the proposed logistic model in Fig. 5.31, which indicates good agreement between the parametric and non-parametric solutions.

Nonparametric methods of regression other than a kernel approach are also possible. Hastie (1983) has presented a scheme using localized logistic regression in analyzing binary data, which involves fitting a logistic model at various locations $\underline{x}$, using each time only a subset of data in the neighborhood of $\underline{x}$. This method is intermediate between a kernel approach and a fully parametric formulation, and deserves future consideration as a tool for analyzing liquefaction data.

Table 5.1

Representative Computer Run Times
of Various Kernel Estimates

Item	Approximate CPU Time
One Calculation of CVE or CVL:	
Uniform Size Kernel	2 sec.
Variable Size Kernel	3 sec.
Integrated Kernel	2 min.
Optimization of CVE or CVL using IMSL Subroutine ZXMIN:	
2 Parameters: Uniform Product Kernel	2-1/2 min.
2 Parameters: Variable Kernel . . . . . (other parameters fixed)	4 min.
2 Parameters: Integrated Kernel . . . . . (other parameters fixed)	1-1/2 hr.

Note: CPU times are reported for runs on a Digital VAX 11/782 computer. Calculations are for 182 data points using a 2-dimensional kernel.

Table 5.2

Comparison of Likelihood Statistics
For Various Smoothing Spaces -
Uniform Product Kernels

	Optimized Kernel				
Smoothed Variables	h	γ	CVL	L	Logistic Regression $L(\hat{\beta})$
$u = (N_1)_{60}$	2.5	0.013	-56.60	-44.77	-47.06
$v = CSRN$					
$u = \log_{10}[(N_1)_{60}]$	0.11	0.22	-55.88	-43.64	-55.48
$v = CSRN$					
$u = (N_1)_{60}$	2.4	0.032	-52.57	-42.68	-44.71
$v = \log_{10}(CSRN)$					
$u = \log_{10}[(N_1)_{60}]$	0.087	0.87	-52.37	-42.72	-45.49
$v = \log_{10}(CSRN)$					

Table 5.3

Summary of Kernel Estimation Parameters
Obtained Using Hierarchical Optimization

Smoothing Space: $u = (N_1)_{60}$ $v = \log(CSRN)$

	UNIFORM	VARIABLE	INTEGRATED VARIABLE
PRODUCT	UP h = 2.4 γ = 0.032 ρ = 0 CVL = -52.57 L = -42.68	VP k = 13 b = 0.080 c = 1.4 CVL = -50.36 L = -40.06	IVP k = 13 b = 0.45 c = 3.6 CVL = -46.89 L = -44.14
ORIENTED	UO h = 2.4 γ = 0.040 ρ = 0.78 CVL = -48.19 L = -39.62	VO Optimum b = 0 No improvement over uniform kernel	IVO Not investigated Due to b = 0 result for kernal

Smoothing Space: $u = (N_1)_{60}$ $v = CSRN$

	UNIFORM	VARIABLE	INTEGRATED VARIABLE
PRODUCT	UP h = 2.5 γ = 0.013 ρ = 0 CVL = -56.58 L = -44.77	VP k = 9 b = 0.79 c = 1.7 CVL = -50.66 L = -40.33	IVP k = 9 b = 0.70 c = 3.5 CVL = -47.17 L = -44.29
ORIENTED	UO h = 2.5 γ = 0.018 ρ = 0.78 CVL = -51.58 L = -42.93	VO k = 6 b = 0.69 c = 2.2 CVL = -49.63 L = -41.06	IVO k = 6 b = 0.83 c = 5.1 CVL = -52.71 L = -50.32

Table 5.4

Kernel Estimation Parameters
for Integrated Uniform Kernel

Smoothing Space: $u = (N_1)_{60}$ $v = \log(CSRN)$

	INTEGRATED UNIFORM
PRODUCT	IUP h = 4.9 γ = 0.029 ρ = 0 CVL = -47.24 L = -44.70
ORIENTED	IUO h = 4.9 γ = 0.031 ρ = 0.25 CVL = -46.95 L = -44.13

Smoothing Space: $u = (N_1)_{60}$ $v = CSRN$

	INTEGRATED UNIFORM
PRODUCT	IUP h = 4.9 γ = 0.11 ρ = 0 CVL = -49.47 L = -45.63
ORIENTED	IUO h = 4.9 γ = 0.12 ρ = 0.32 CVL = -49.11 L = -45.46

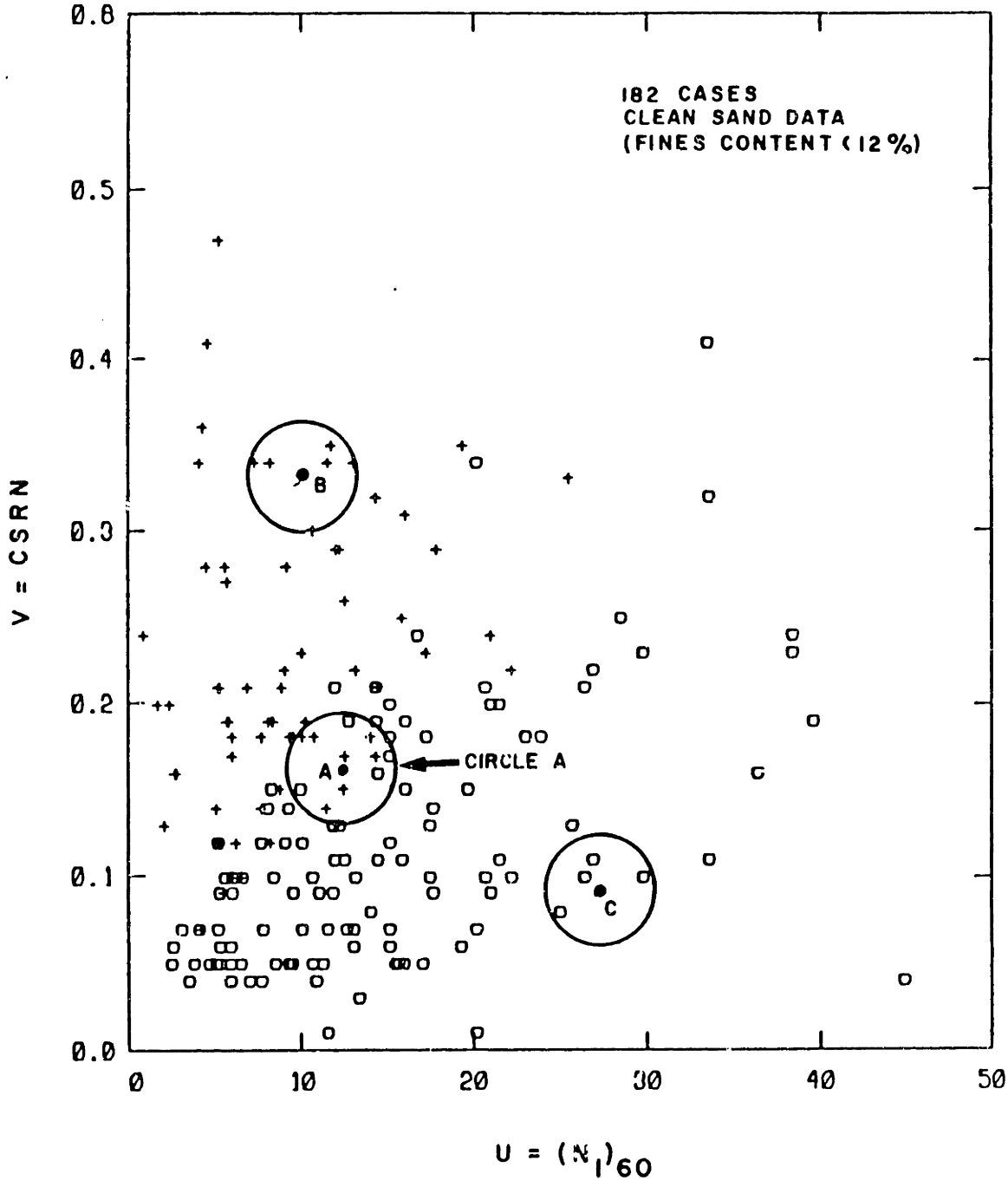
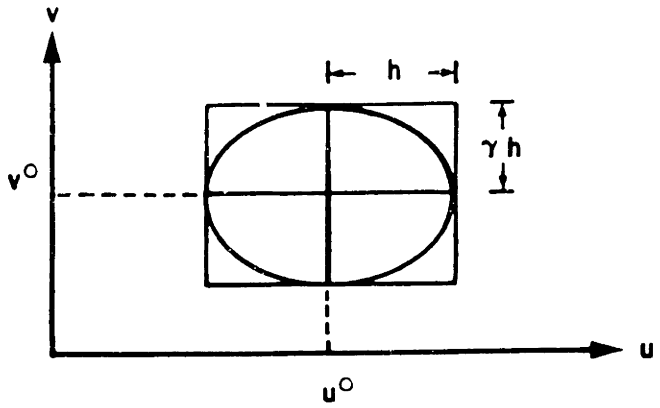
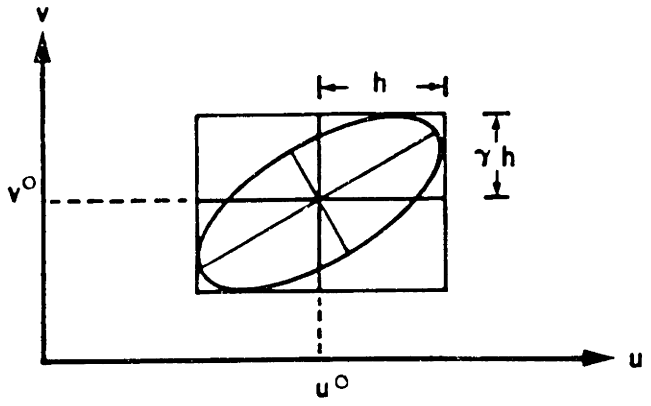


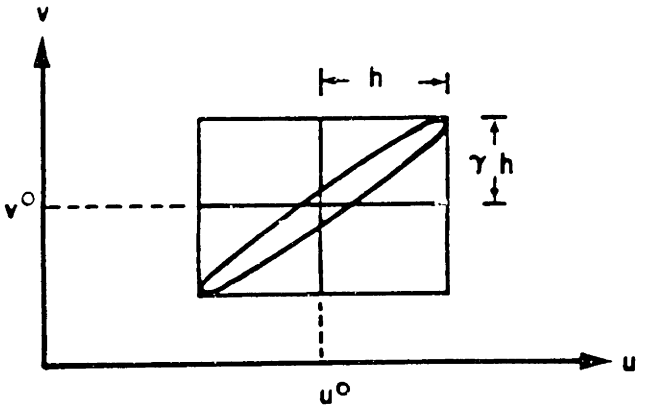
Fig. 5.1 Schematic of the basic concept of kernel estimation.



(a) $\rho = 0$



(b) $\rho > 0$



(c) $\rho \rightarrow 1$

Fig. 5.2 Schematic of the ellipse in a two-dimensional space defining constant values of the statistical (Mahalanobis) distance.

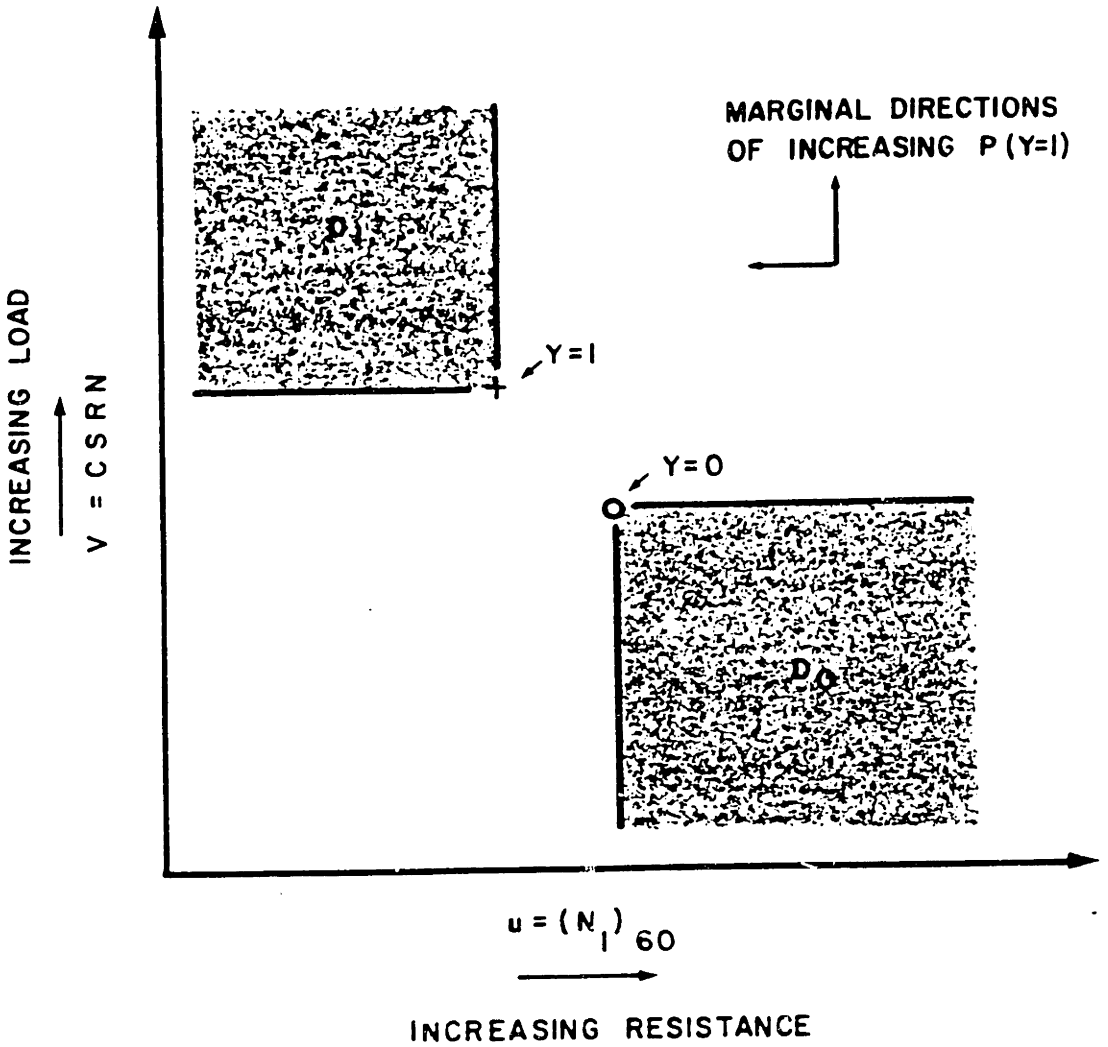


Fig. 5.3 Schematic of the rationale and notation for the integrated kernel.

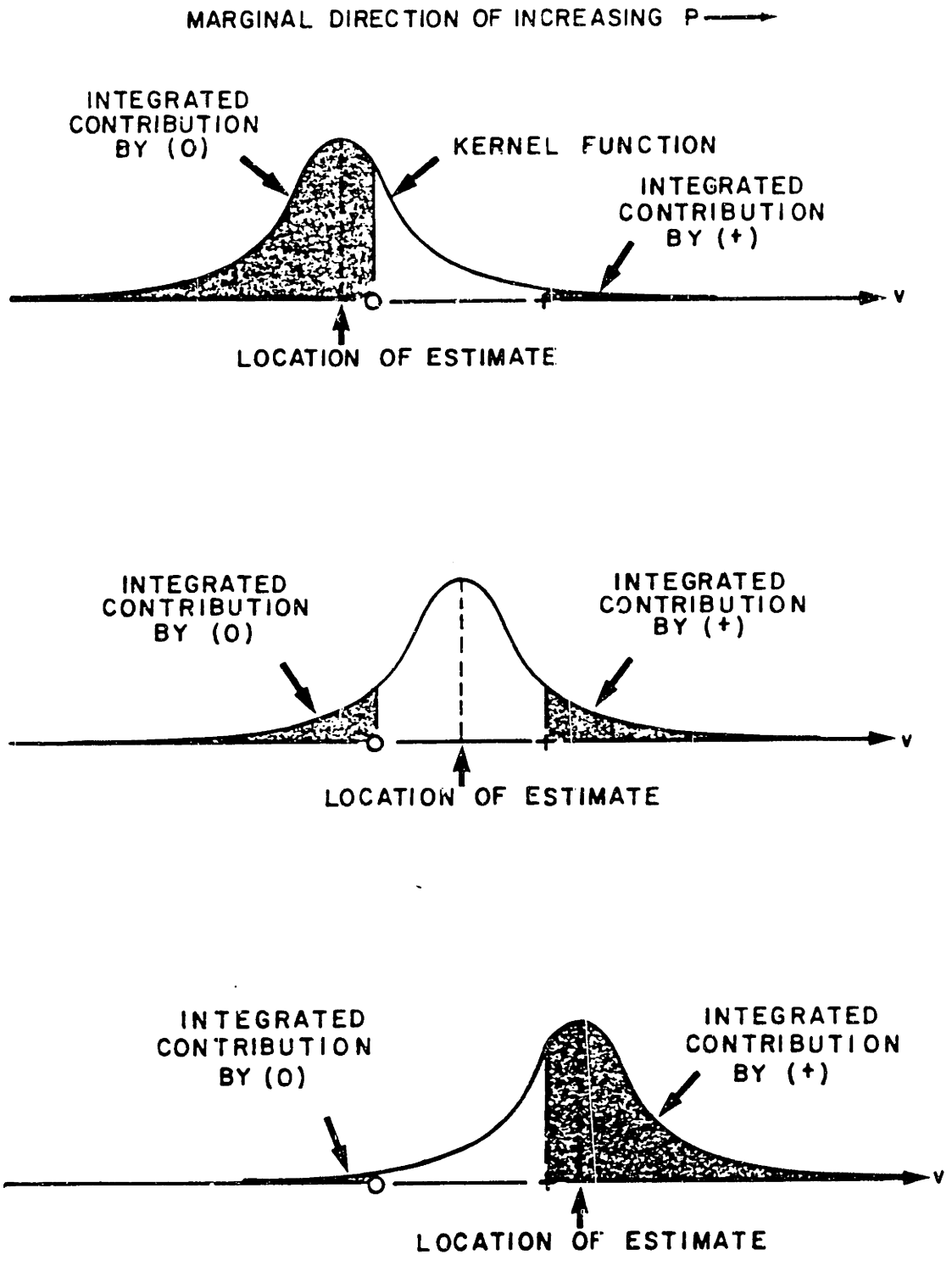
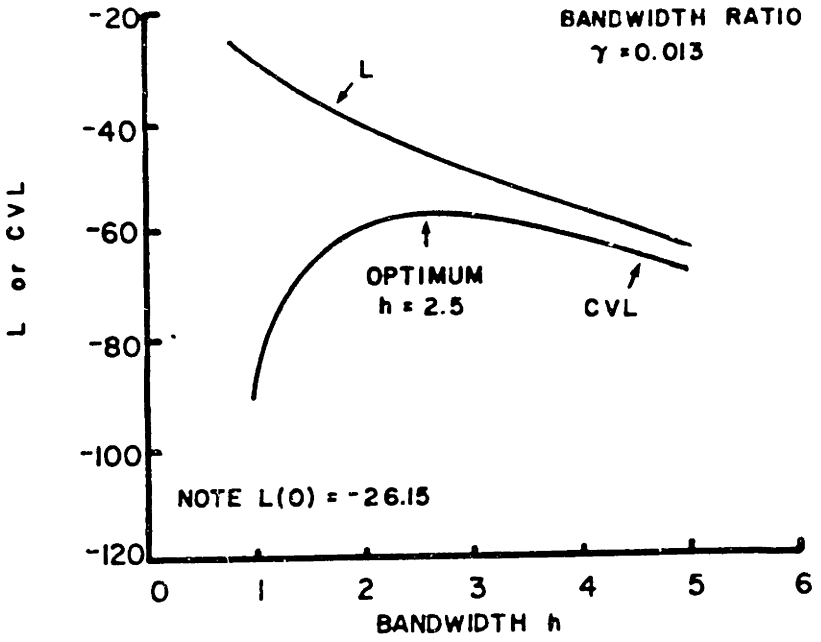
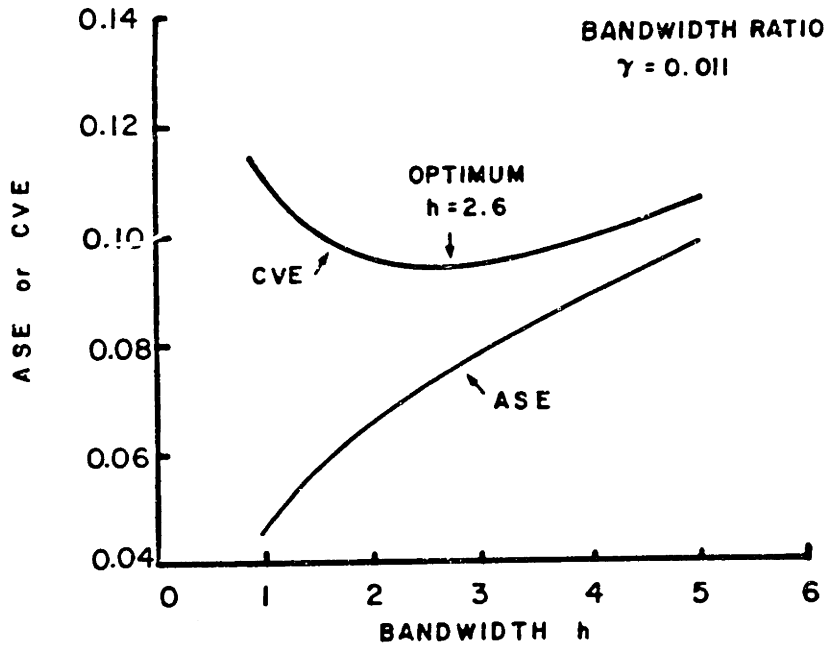


Fig. 5.4 Schematic one-dimensional analog in the marginal v-direction of the integrated kernel smoother.

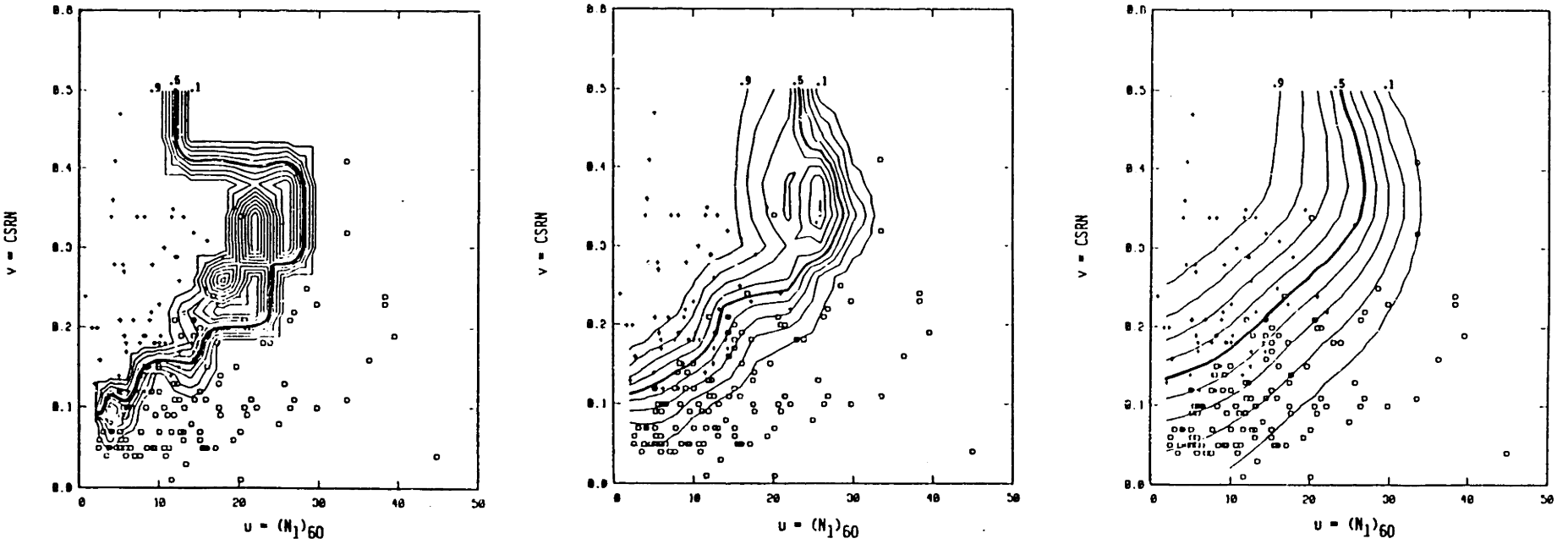


(a) LIKELIHOOD FUNCTION



(b) SQUARED ERROR FUNCTION

Fig. 5.5 Comparison of the likelihood and squared error functions used in cross-validation optimization procedure. Results shown for uniform product kernel in smoothing space u = $(N_1)_{60}$, v = CSRN.

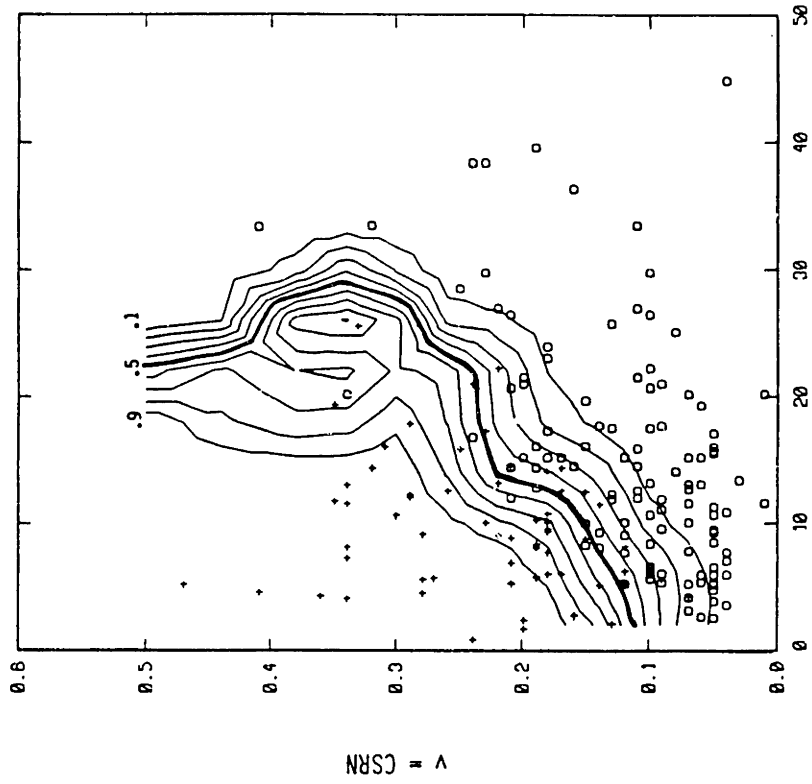


(a) Bandwidth too small (h = 0.5)

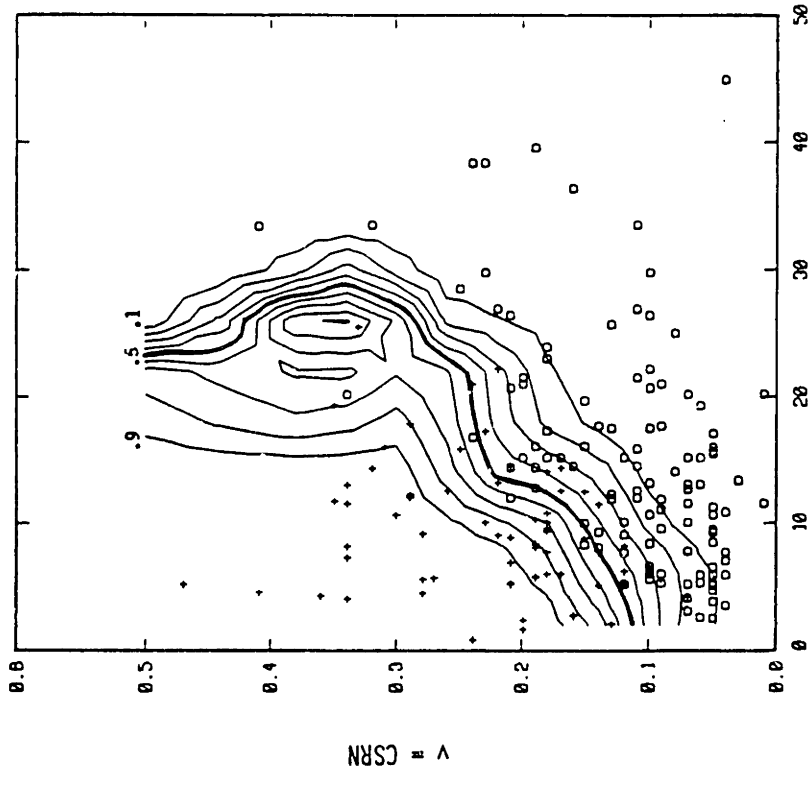
(b) Optimal bandwidth (h = 2.5)

(c) Bandwidth too large (h = 5.0)

Fig. 5.6 Comparison of effects of varying bandwidths h on estimates of P(Y=1) using a product kernel with $\gamma = 0.013$. Contour lines indicate equal values of $\hat{P}$.



(b) Minimum CVE estimate
h = 2.6, γ = 0.011, ρ = 0



(a) Maximum CVL estimate
h = 2.5, γ = 0.013, ρ = 0

Fig. 5.7 Comparison of a maximum CVL and minimum CVE estimates $\hat{P}(Y=1)$.
Contour lines indicate equal values of $\hat{P}$.

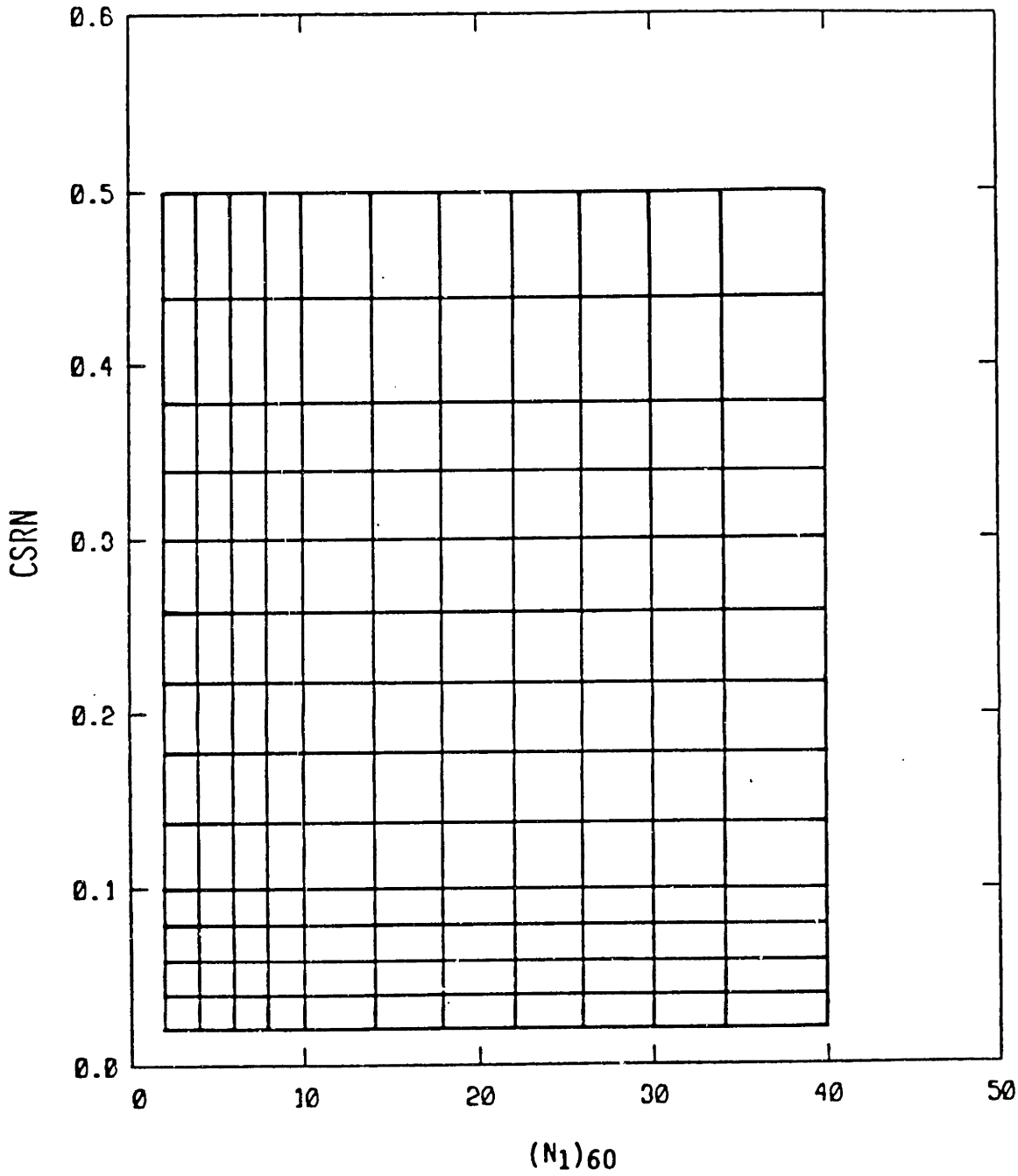


Fig. 5.8 Grid of points at which $\hat{P}(Y=1)$ are calculated for the analysis of liquefaction data. Linear interpolation used between grid points to draw contour lines.

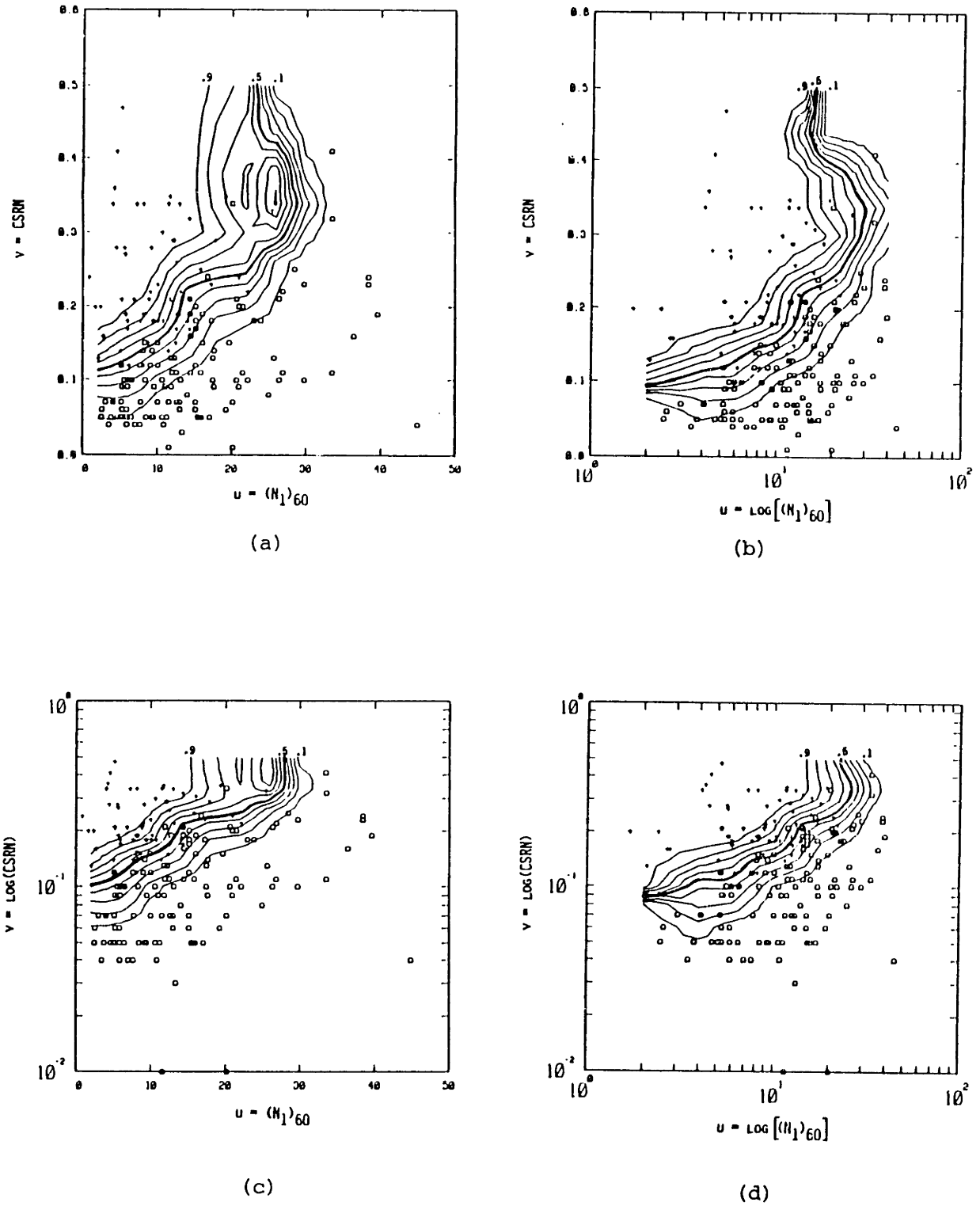
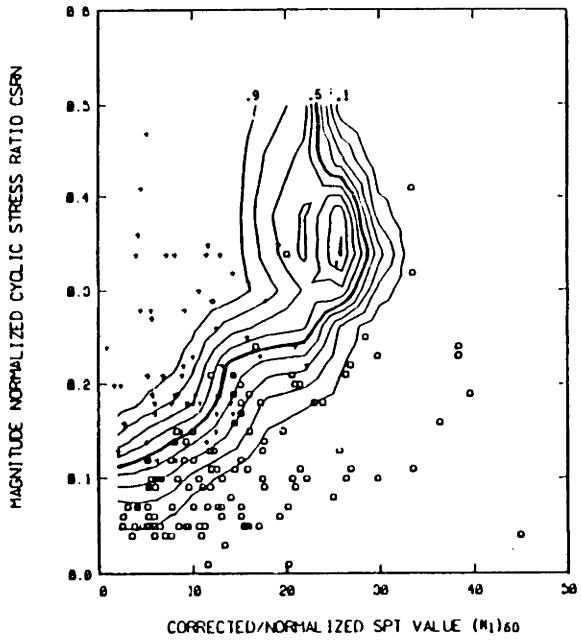
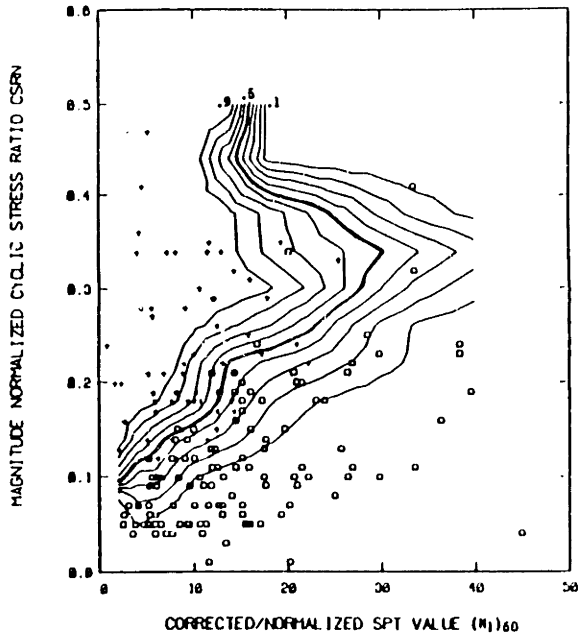


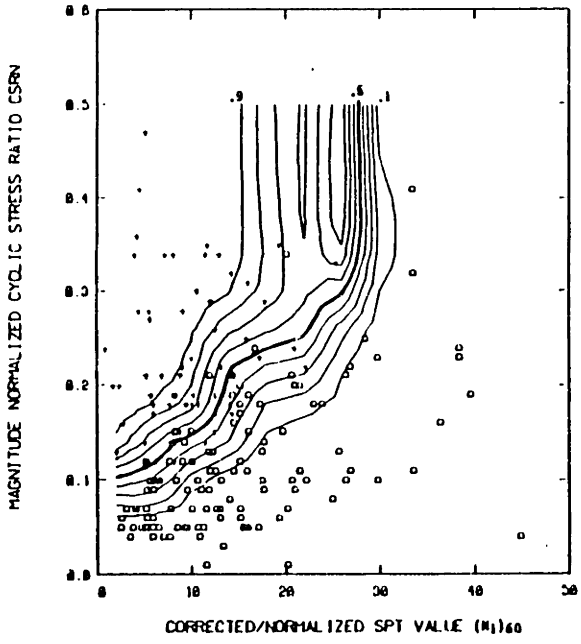
Fig. 5.9 Comparison of $\hat{P}(Y=1)$ estimates in various smoothing spaces (transformed axes) - uniform product kernels. Contour lines indicate equal values of $\hat{P}$.



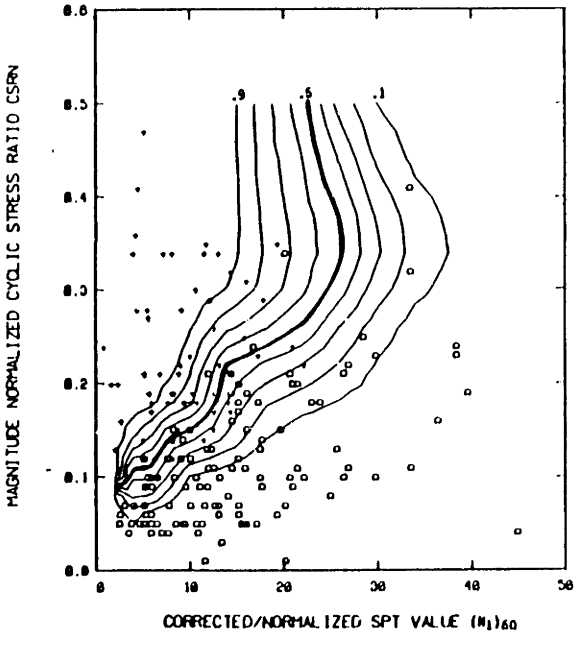
(a) $u = (N_1)_{60}$
$v = CSRN$



(b) $u = \log[(N_1)_{60}]$
$v = CSRN$

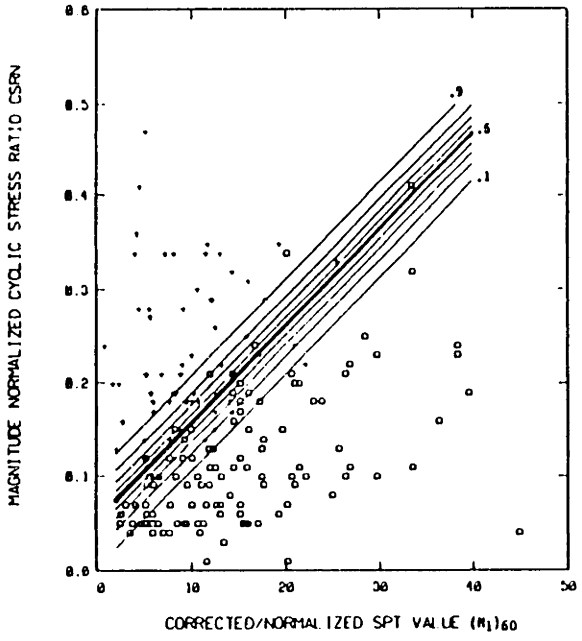


(c) $u = (N_1)_{60}$
$v = \log(CSRN)$

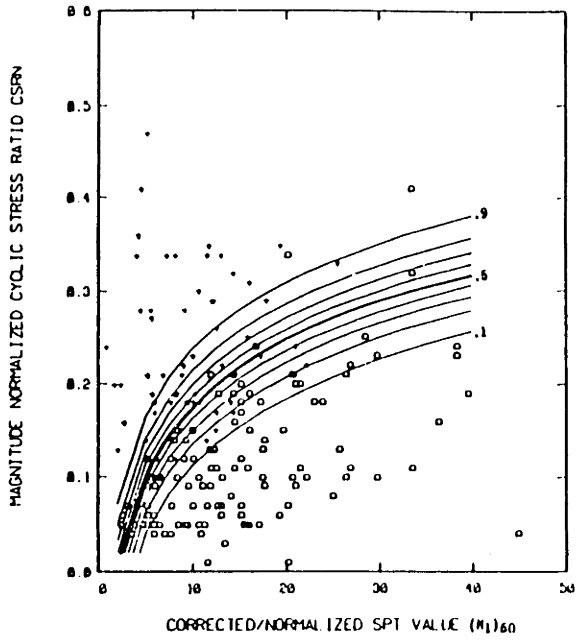


(d) $u = \log[(N_1)_{60}]$
$v = \log(CSRN)$

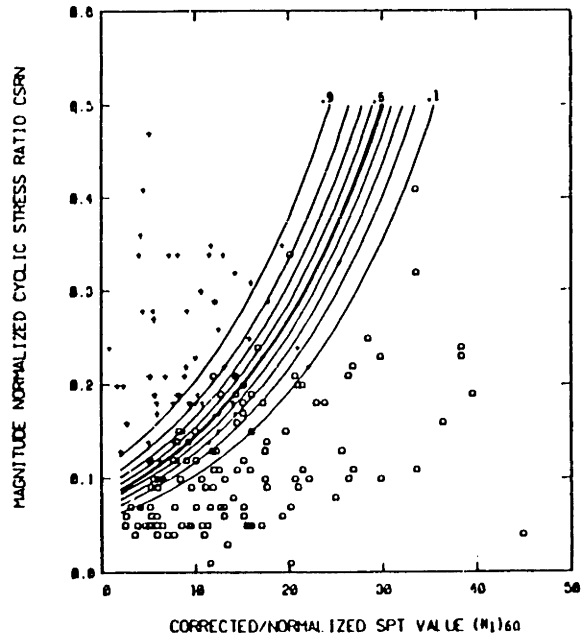
Fig. 5.10 Comparison of $\hat{P}(Y=1)$ estimates in various smoothing spaces (natural axes) - uniform product kernels. Contours indicate equal values of $\hat{P}$.



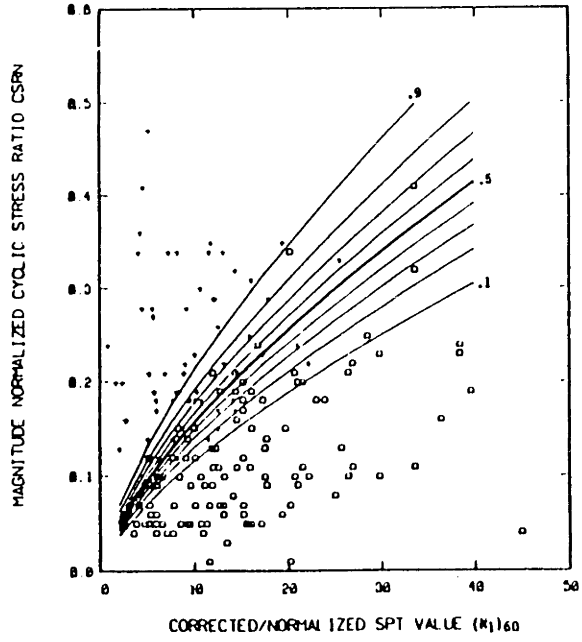
(a) $(N_1)_{60}$ - CSRN fit



(b) $\ln[(N_1)_{60}]$ - CSRN fit

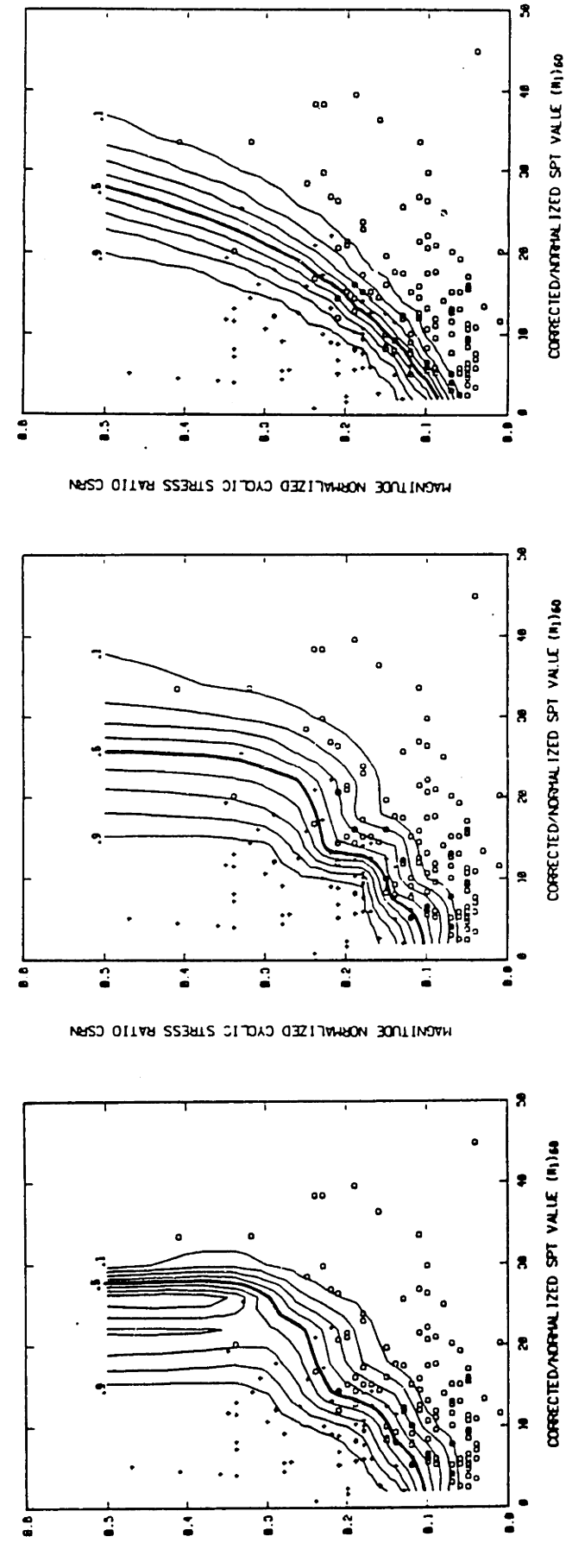


(c) $(N_1)_{60}$ - ln(CSRN) fit



(d) $\ln[(N_1)_{60}]$ - ln(CSRN) fit

Fig. 5.11 Equivalent parametric results to those in Fig. 5.10 obtained using logistic regression. Contour lines indicate equal values of $\hat{P}(Y=1)$.

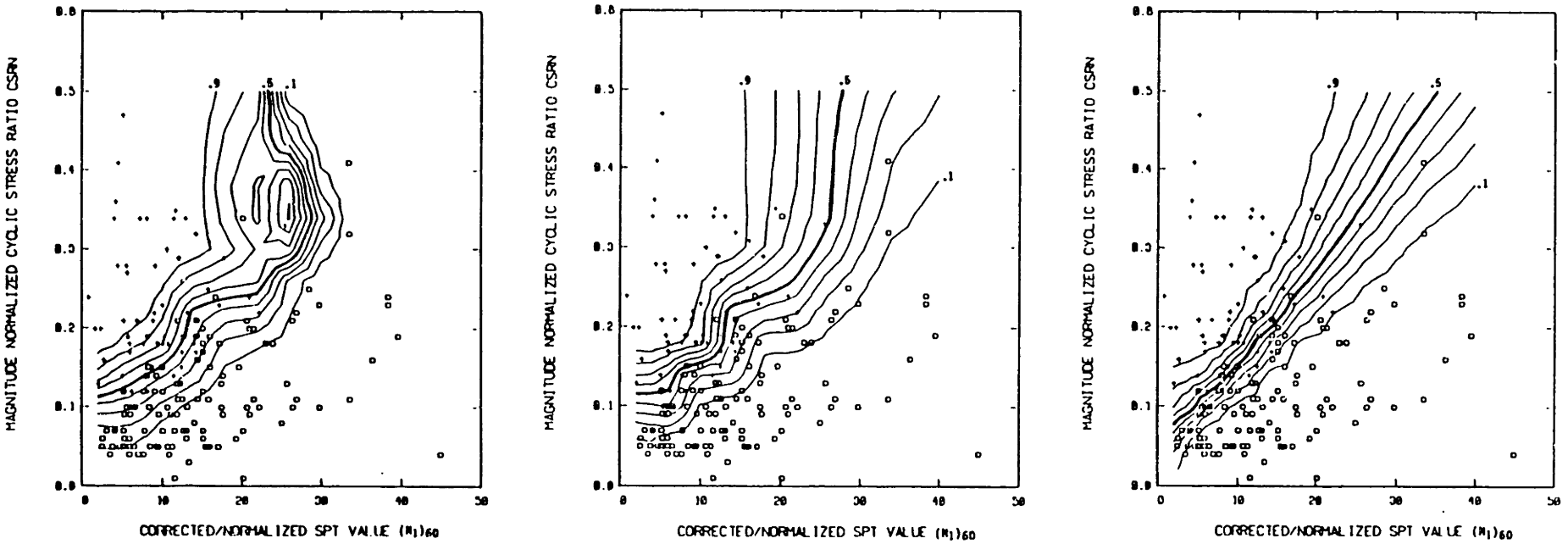


(a) Uniform (UP) kernel

(b) Variable (VP) kernel

(c) Integrated Variable (IVP) Kernel

Fig. 5.12 Estimates $\hat{P}(Y=1)$ using uniform, variable, and integrated variable product kernels. Smoothing space: $u = (N_1)60$, $v = \log(CSRN)$. Contour lines indicate equal values of $\hat{P}$.



(a) Uniform (UP) kernel

(b) Variable (VP) kernel

(c) Integrated variable (IVP) kernel

Fig. 5.13 Estimates $\hat{P}(Y=1)$ using uniform, variable and integrated, variable product kernels. Smoothing space: $u = (N_1)_{60}$, $v =$ CSRN. Contour lines indicate equal values of $\hat{P}$.

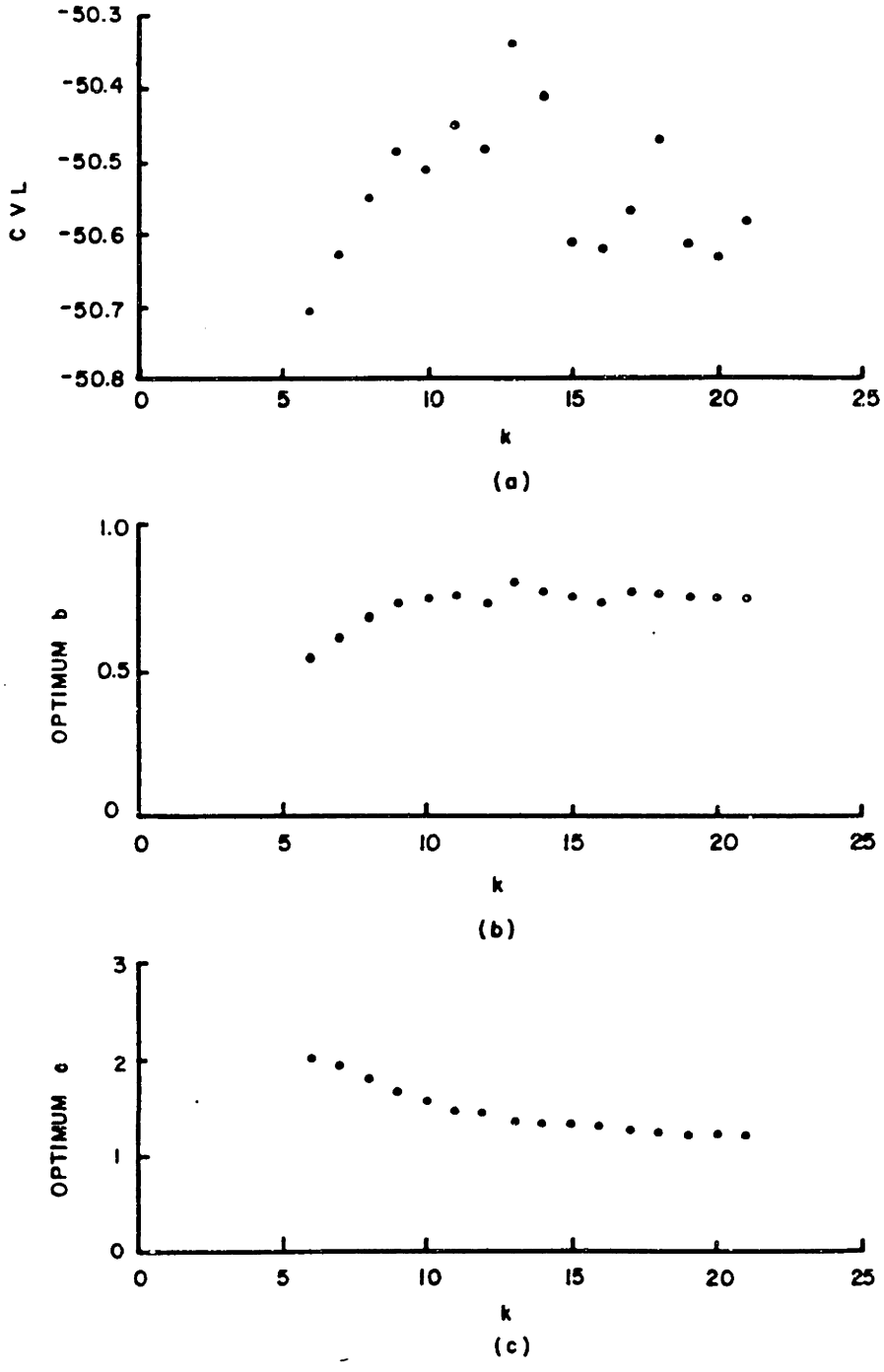
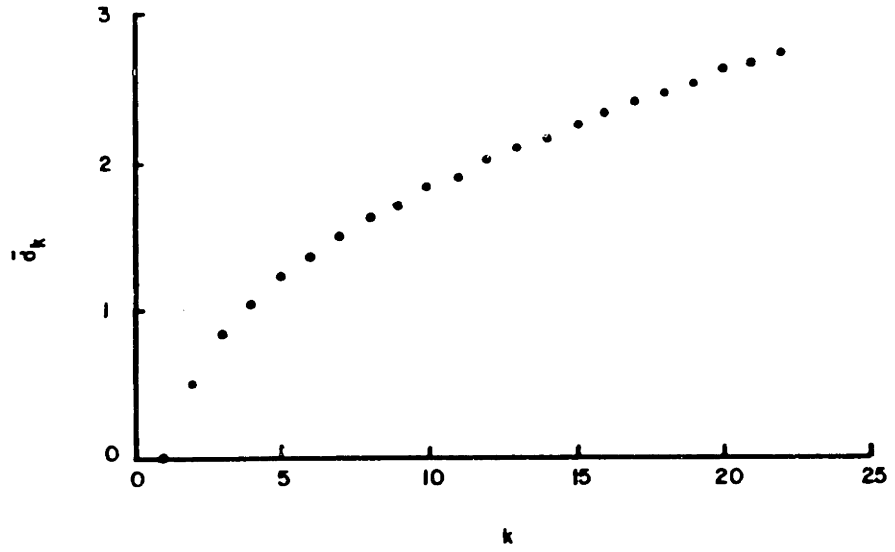
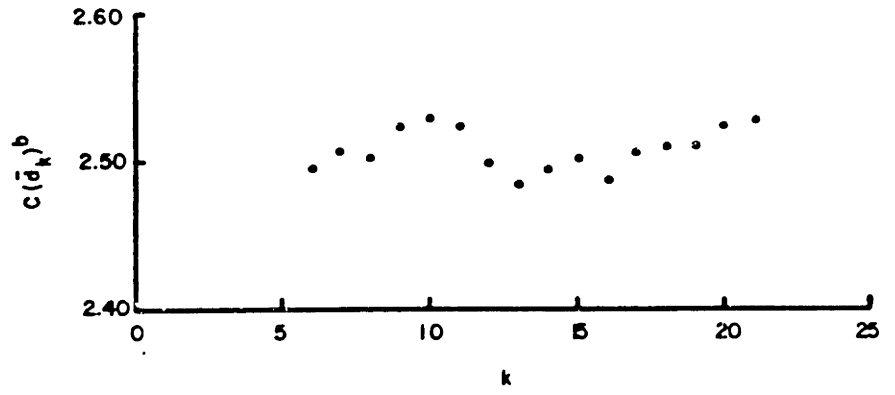


Fig. 5.14 CVL and optimum values of b and c for a variable product kernel. Smoothing space: $u = (N_1)_{60}$, $v = \log(\text{CSRN})$.

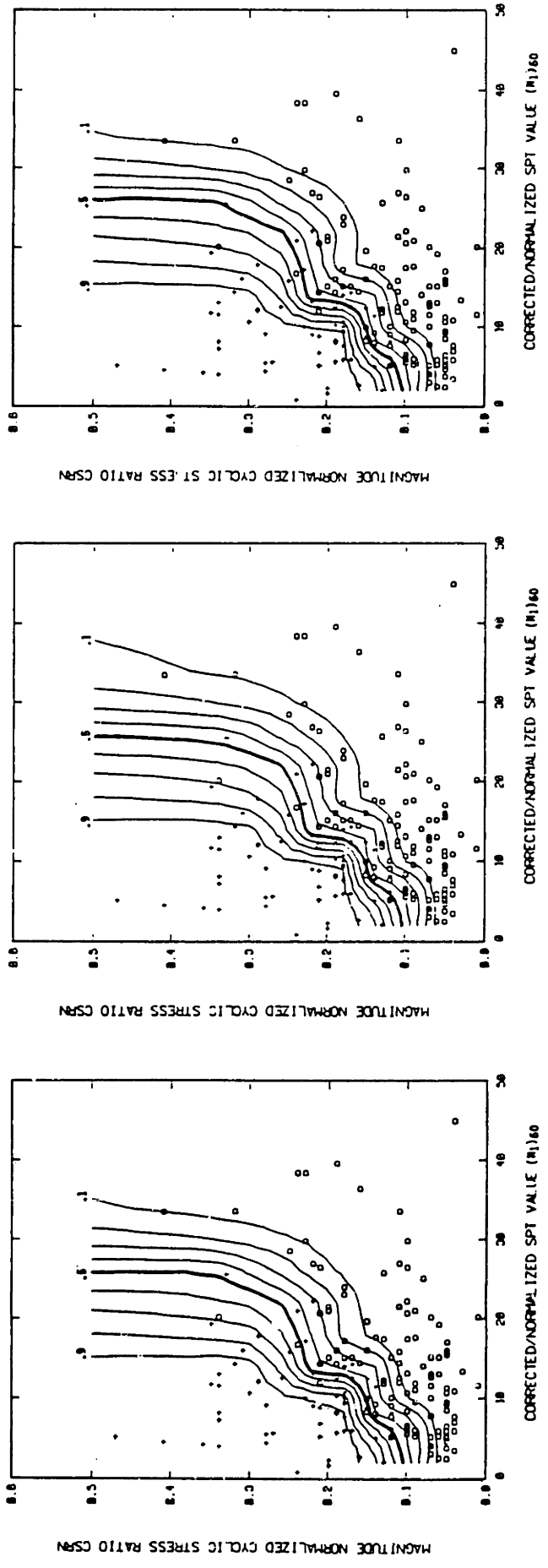


(a)



(b)

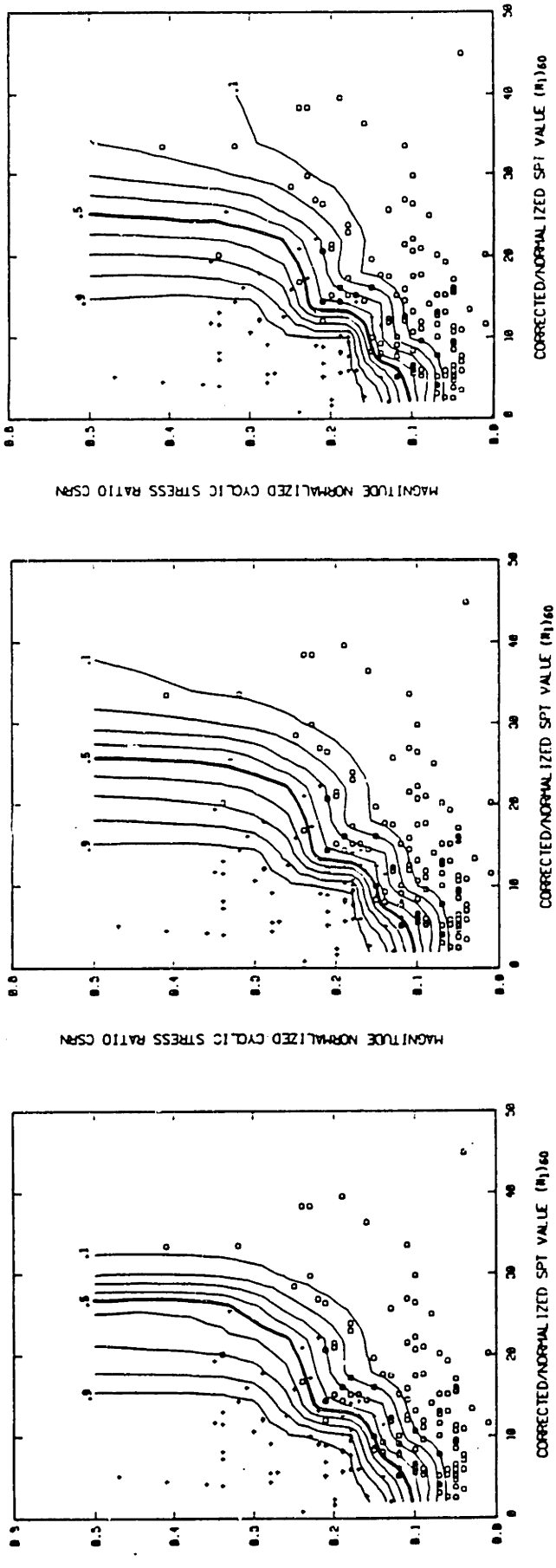
Fig. 5.15 Mean distance $\bar{d}_k$ to the kth nearest neighbor and $c(\bar{d}_k)^b$ as a function of k. Parameters b and c are optimal parameters for each value of k. Smoothing space: $u = (N_1)_{60}$, $v = \log(CSRN)$.



(a) k = 9 (b) k = 13 (c) k = 16

(optimum)

Fig. 5.16 Comparison of estimates $\hat{P}(Y=1)$ using variable product kernels with different k values. Parameters b and c were optimized assuming a fixed k. Smoothing space: $u = (N_1)60$, $v = \log(CSRN)$. Contour lines indicate equal values of $\hat{P}$.



(a) b = 0.50

(b) b = 0.80
(optimum)

(c) b = 1.00

Fig. 5.17 Comparison of estimates $\hat{P}(Y=1)$ using variable product kernels with different b values for k = 13. Paramter c was optimized assuming fixed k and b. Smoothing space: $u = (N_1)_{60}$, $v = \log(CSRN)$. Contour lines indicate equal values of $\hat{P}$.

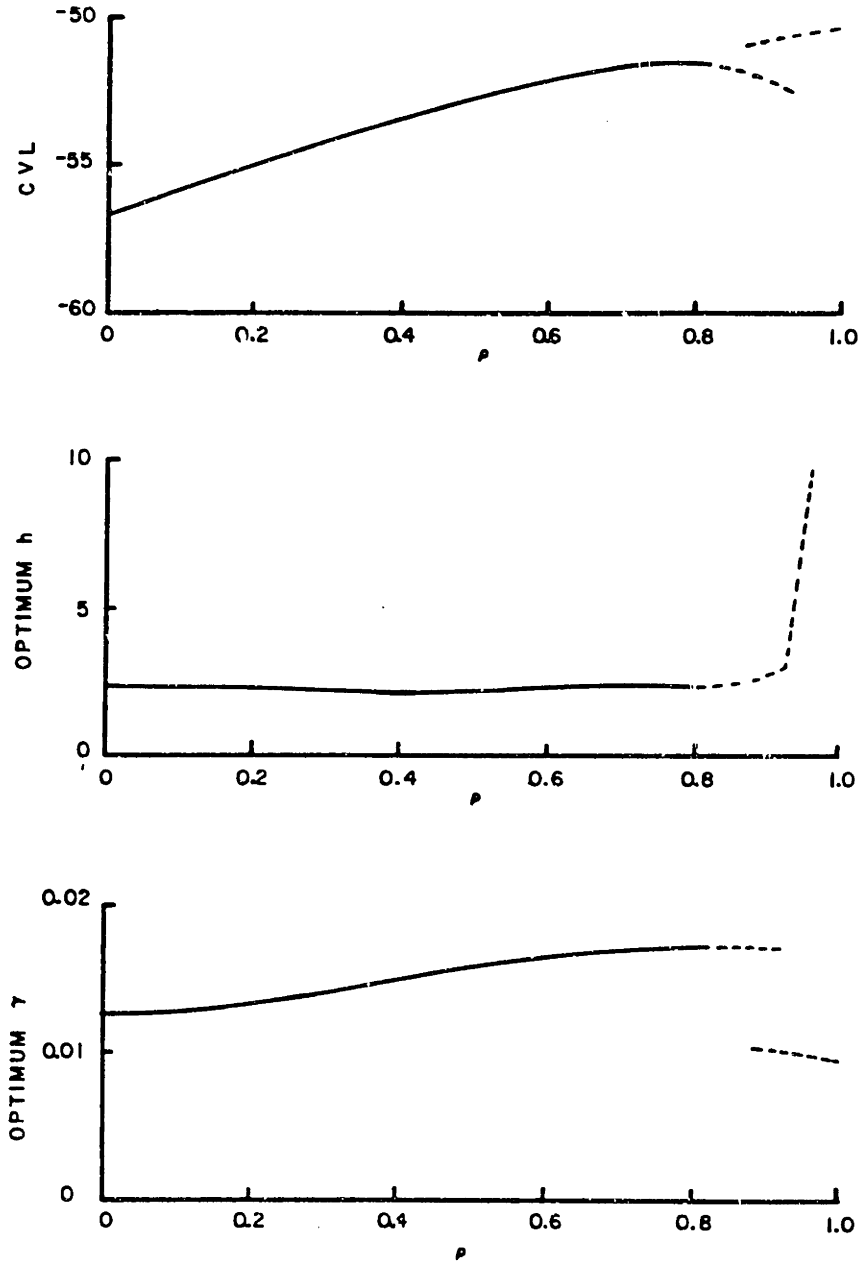


Fig. 5.18 CVL and optimum values of h and γ as a function of ρ for a uniform kernel. Smoothing space: $u = (N_1)_{60}$, v = CSRN.

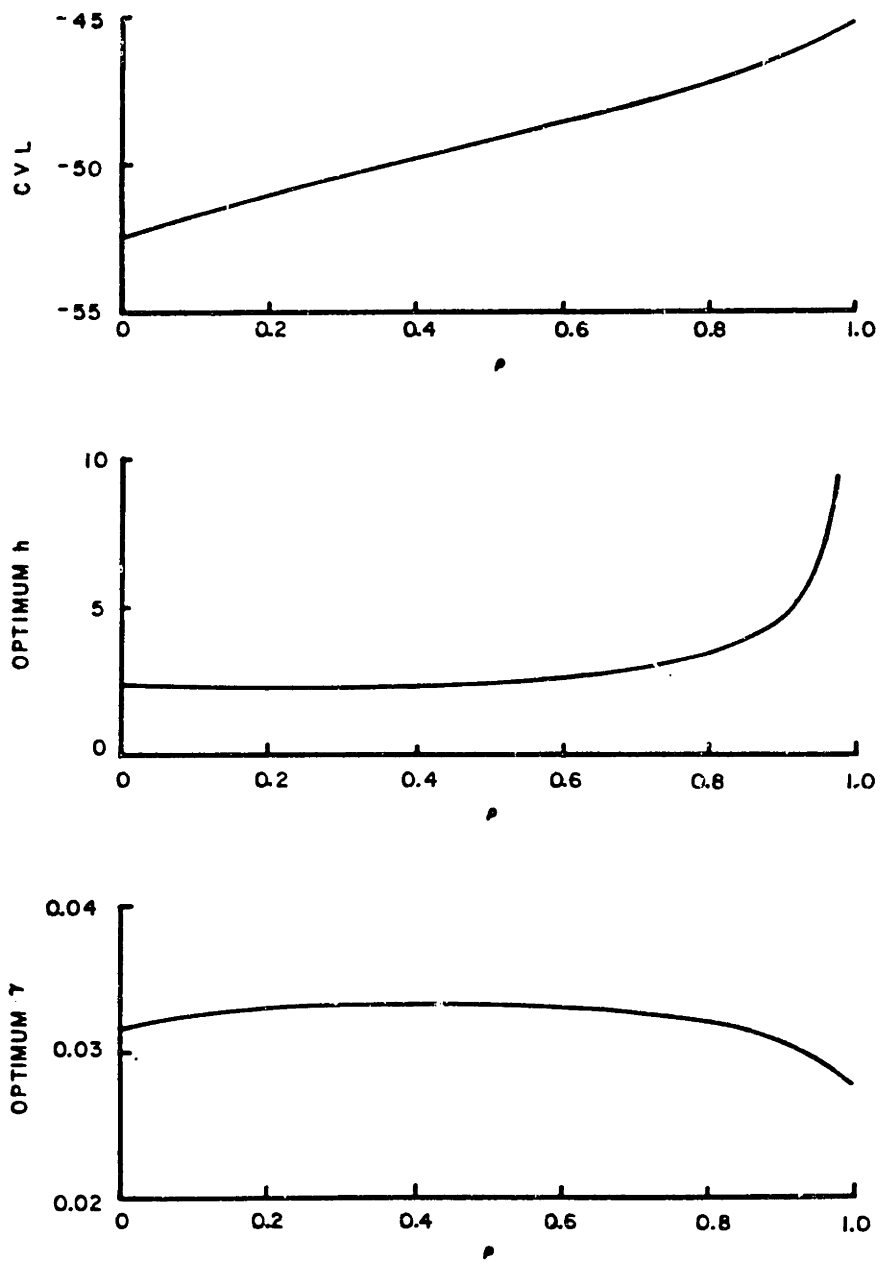
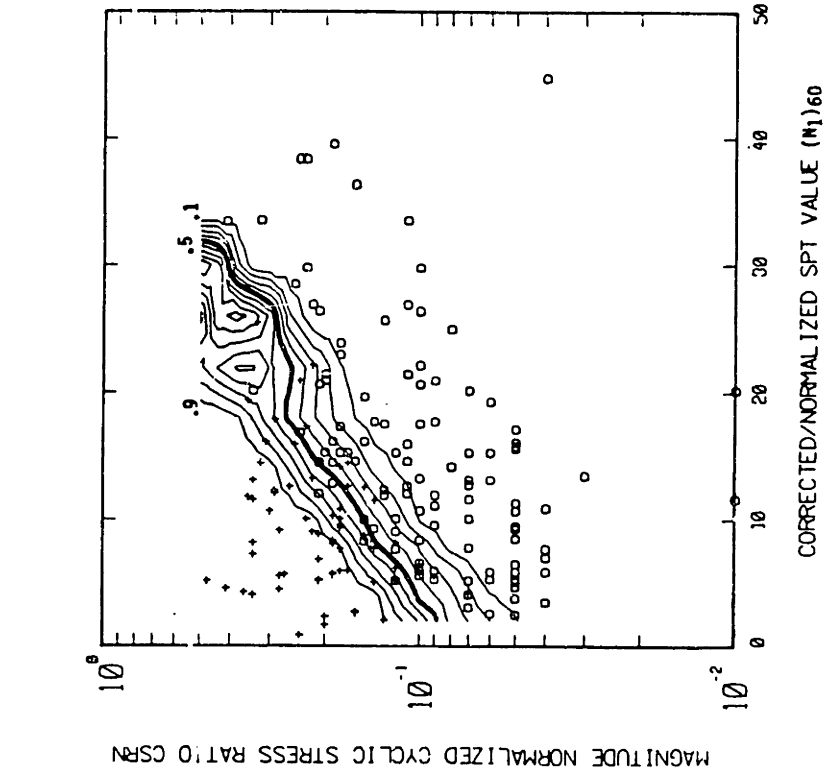


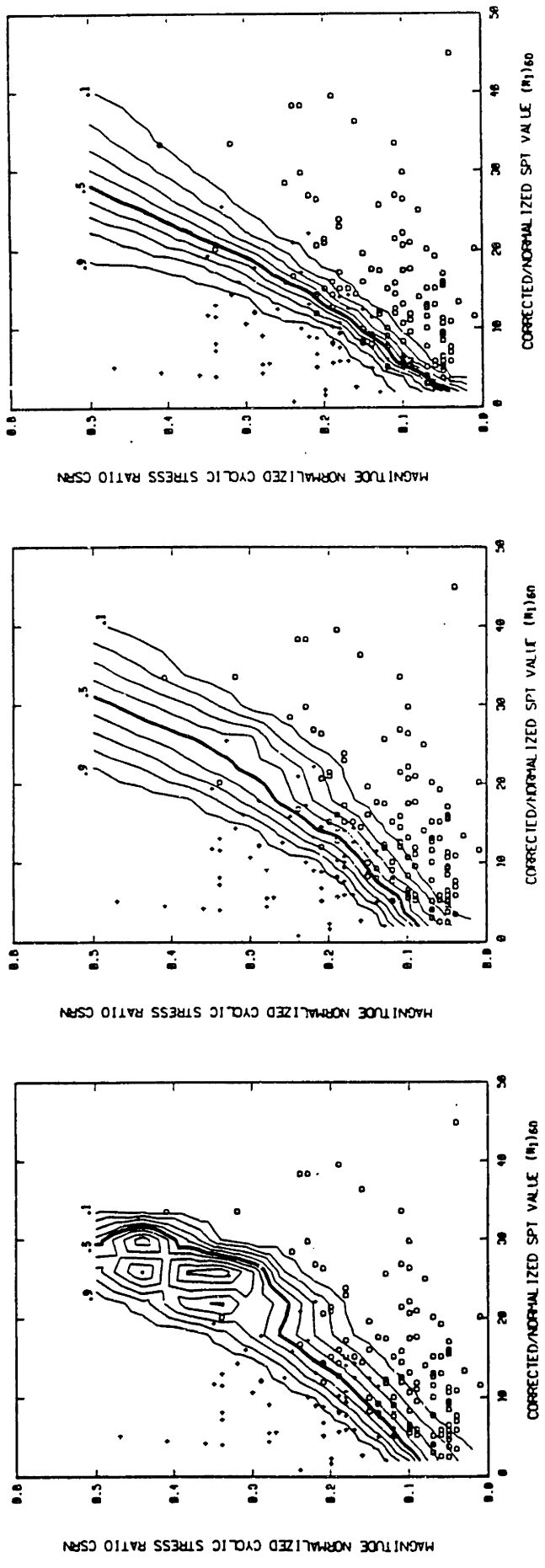
Fig. 5.19 CVL and optimum values of h and γ as a function of ρ for a uniform kernel. Smoothing space: $u = (N_1)_{60}$, $v = \log(\text{CSRN})$.



(a) Smoothing space:
$u = (N_1)60$, $v = CSRN$

(b) Smoothing space
$u = (N_1)60$, $v = \log(CSRN)$

Fig. 5.20 Comparison of two different oriented kernel $\hat{P}$ estimates in different smoothing spaces. Note the tendency toward straight line contours of $\hat{P}$ in both spaces, but that the straight lines are more appropriate for the smoothing space shown in (b).

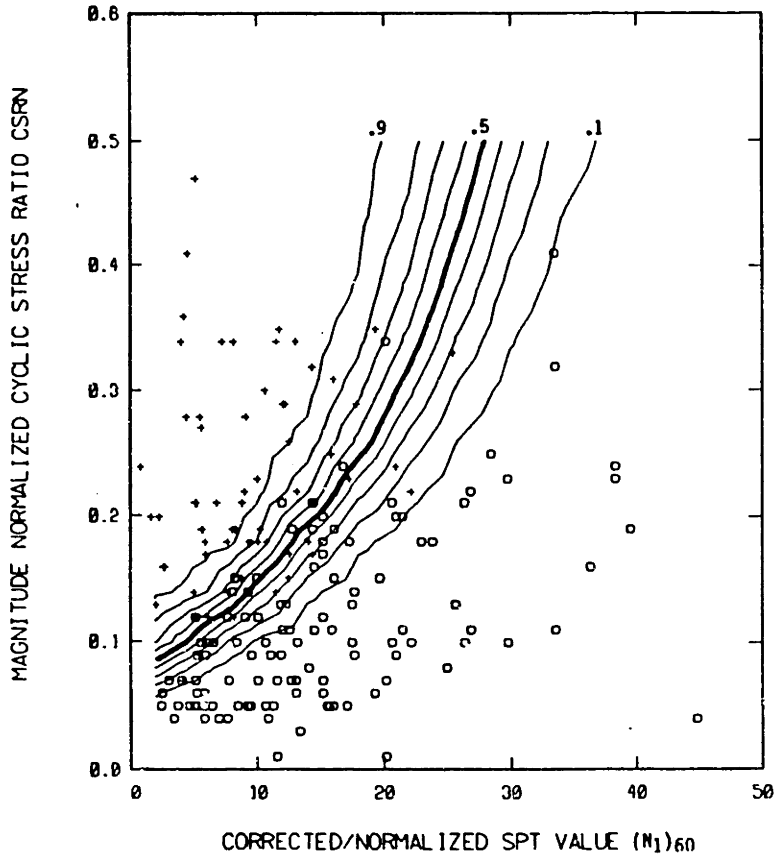


(a) Uniform (UO) Kernel

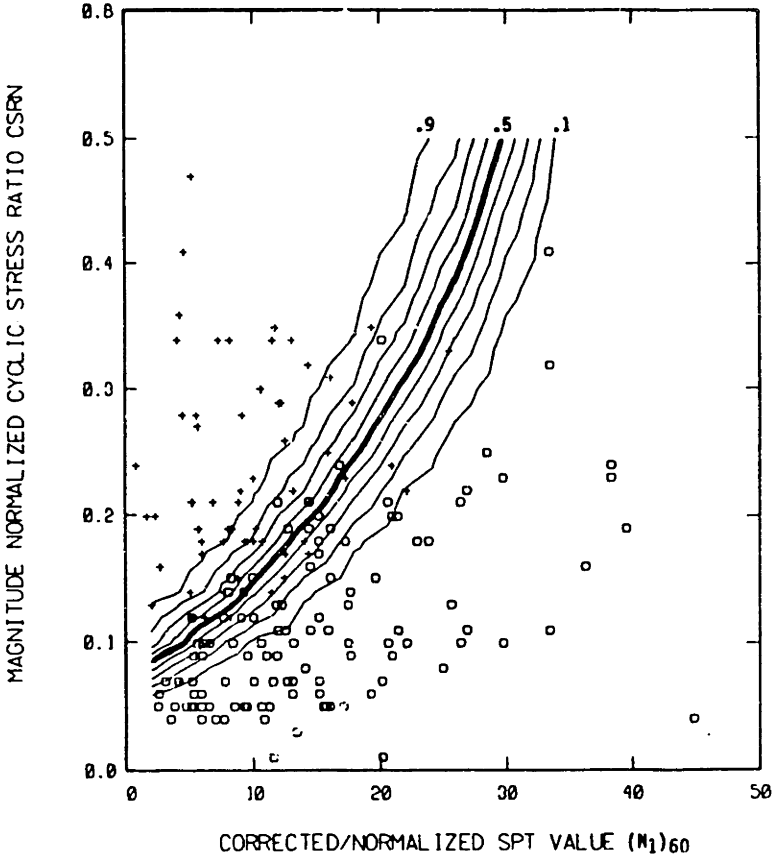
(b) Variable (VO) Kernel

(c) Integrated variable (IVO) kernel

Fig. 5.21 Estimates $\hat{P}(Y=1)$ using uniform, variable, and integrated variable oriented kernels. Smoothing space: $u = (N_1)60$, $v = CSRN$. Contour lines indicate equal values of $\hat{P}$. (Note: Compare these results with Fig. 5.13.)

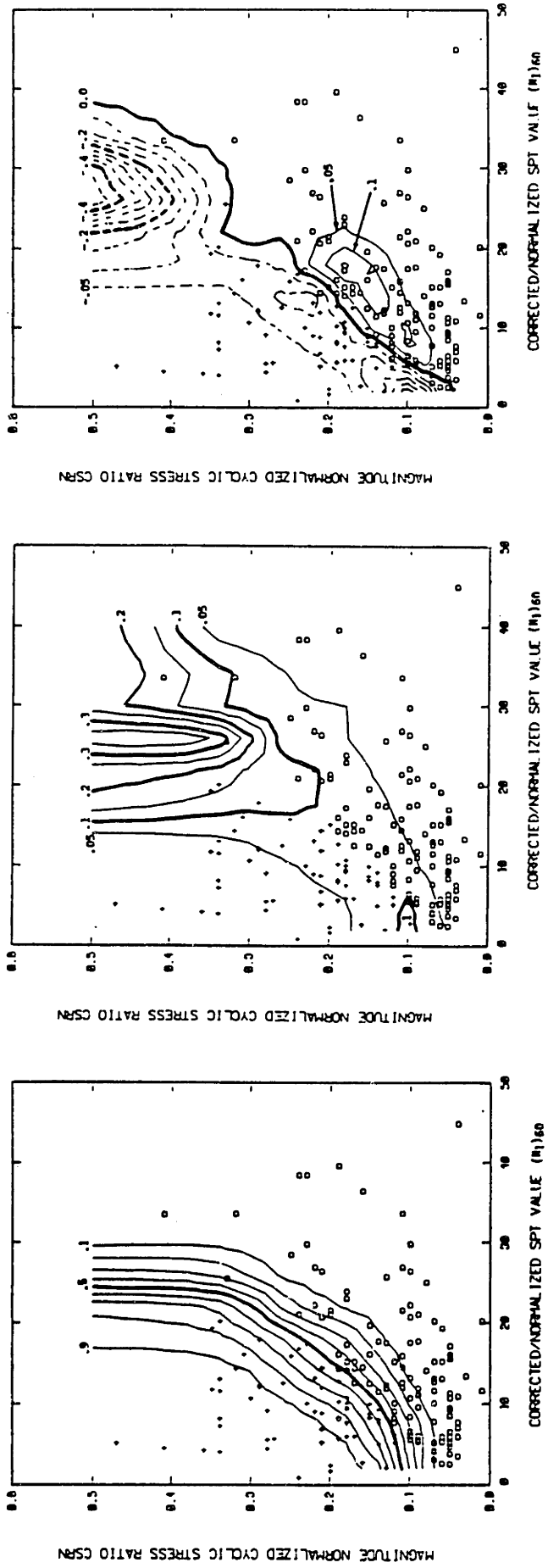


(a) Integrated Variable Product (IVP) kernel



(b) Integrated Uniform Oriented (IUO) kernel

Fig. 5.22 Comparison of $\hat{P}$ estimates using an integrated variable product (IVP) kernel and an integrated uniform oriented (IVO) kernel, with comparable log-likelihood statistics. Smoothing space: $u = (N_1)_{60}$, $v = \log(CSRN)$.



(a) $E[\hat{P}]$ (b) $SD[\hat{P}]$ (c) Bias $[\hat{P}]$

Fig. 5.23 Expected value, standard deviation, and bias of the estimator $\hat{P}(Y=1)$ obtained for a uniform product (UP) kernel. Underlying P assumed to be the best-fit logistic function shown in Fig. 5.11(c). Contour lines indicate equal values of $E[\hat{P}]$, $SD[\hat{P}]$, and Bias $[\hat{P}]$.

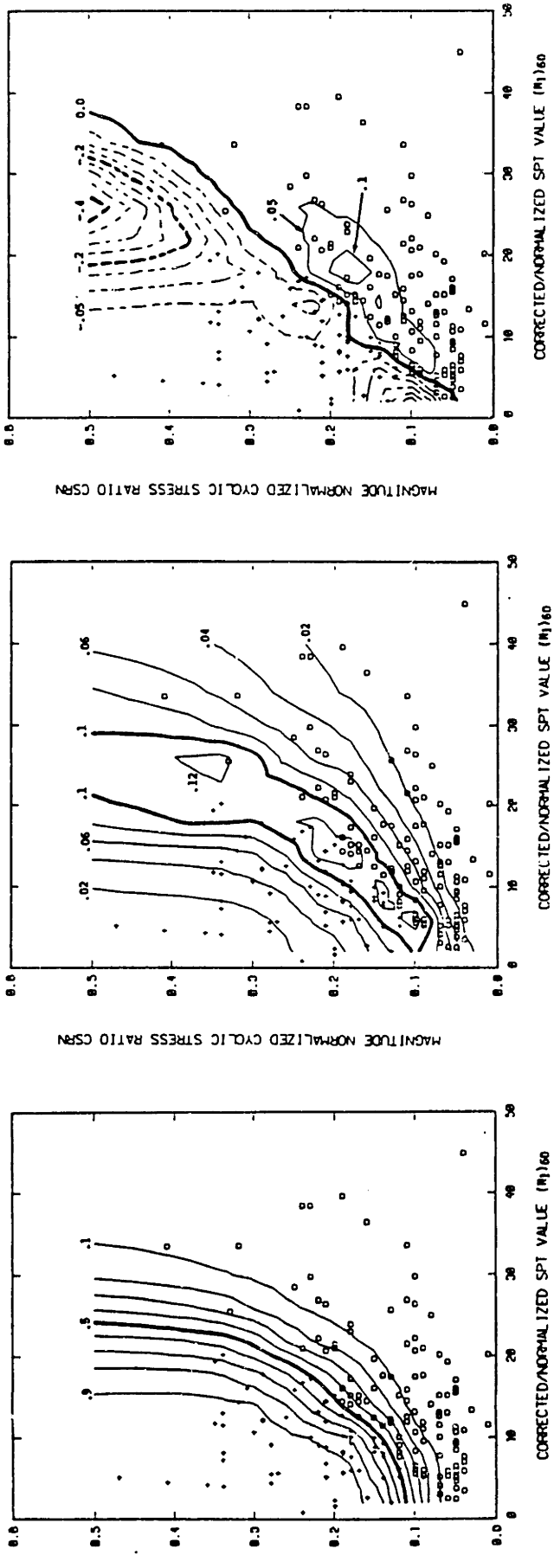
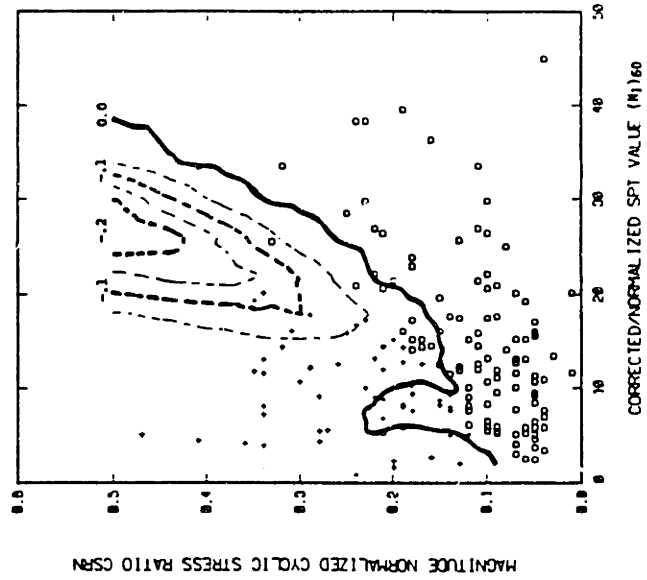
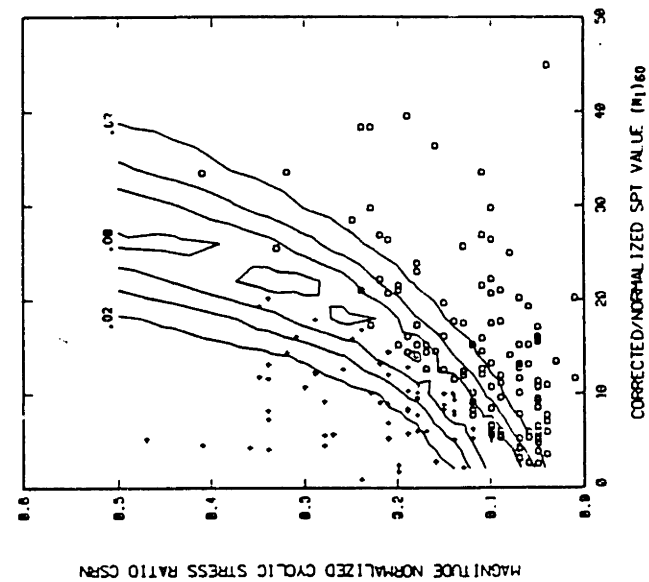


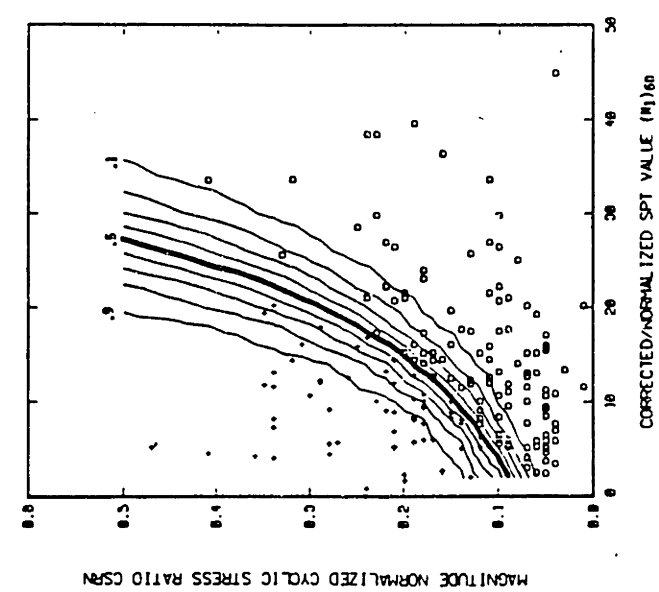
Fig. 5.24 Expected value, standard deviation, and bias on the estimator $\hat{P}(Y=1)$ obtained for a variable product (VP) kernel. Underlying P assumed to be the best-fit logistic function shown in Fig. 5.11(c). Contour lines indicate equal values of $E[\hat{P}]$, $SD[\hat{P}]$, and Bias $[\hat{P}]$.



(c) Bias[$\hat{P}$]

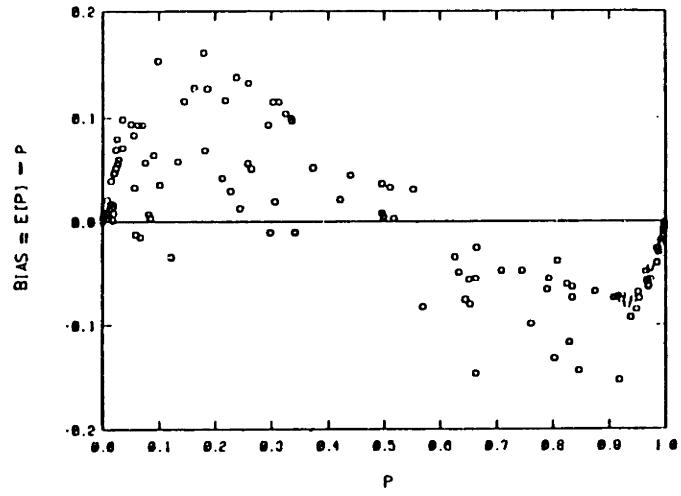


(b) SD[$\hat{P}$]

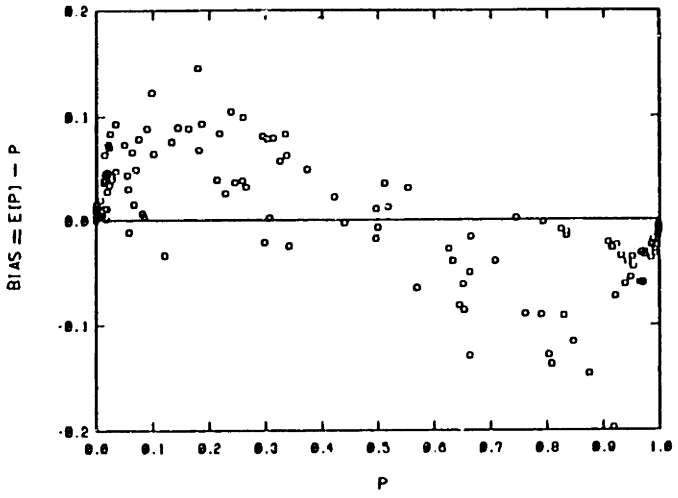


(a) E[$\hat{P}$]

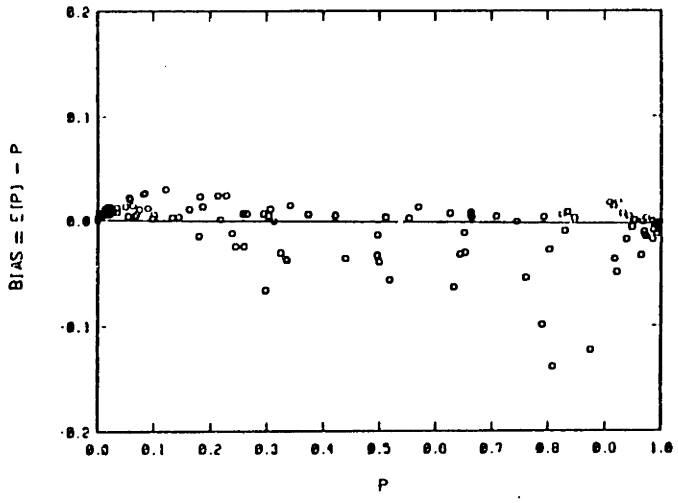
Fig. 5.25 Expected value, standard deviation, and bias of the estimator $\hat{P}(Y=1)$ obtained for an integrated variable product (IVP) kernel. Underlying P assumed to be the best-fit logistic model shown in Fig. 5.11(c). Contour lines indicate equal values of $E[\hat{P}]$, $SD[\hat{P}]$, and Bias $[\hat{P}]$. Note: The results in this figure were obtained through Monte Carlo simulation with number of simulations = 30.



(a) Uniform product (UP) kernel

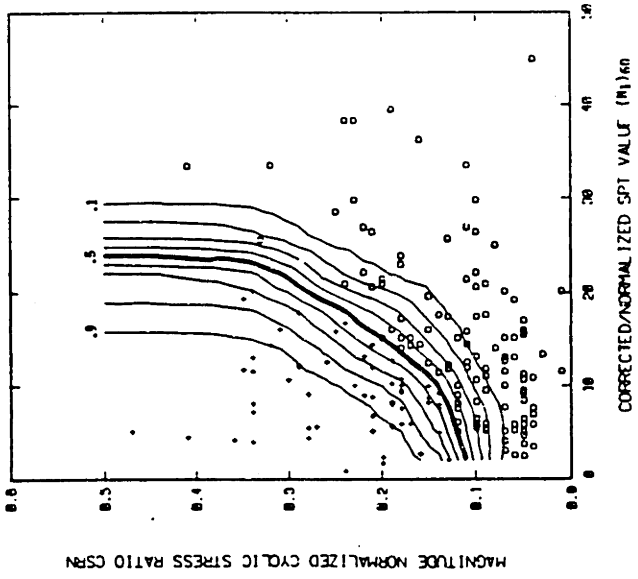


(b) Variable product (VP) kernel

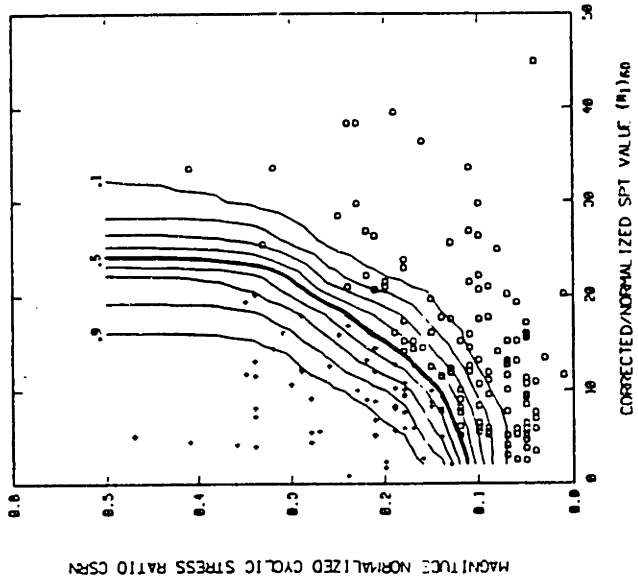


(c) Integrated variable product (IVP) kernel

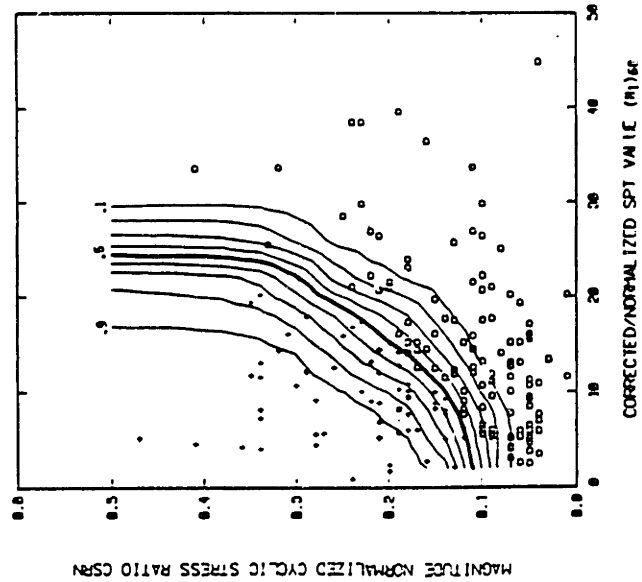
Fig. 5.26 Bias versus "actual" P for UP, VP and IVP kernels. Biases shown are evaluated at the data points.



(c) Theory (No. of simulations → ∞)

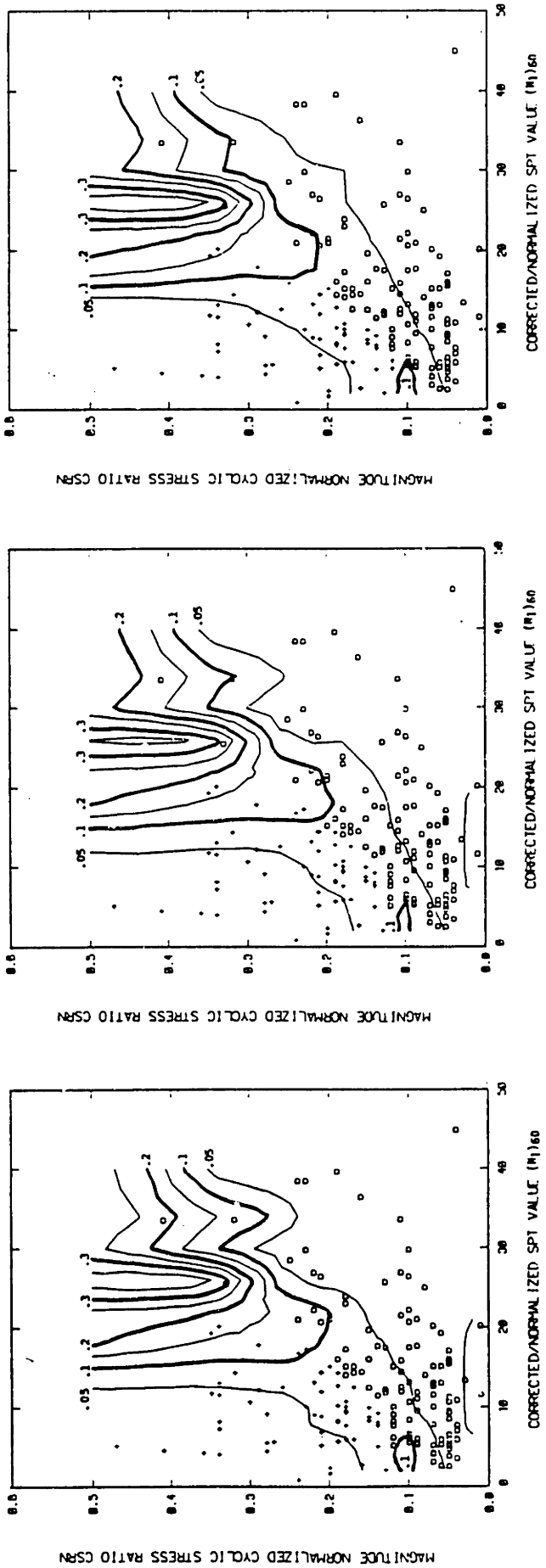


(b) No. of simulations = 60



(a) No. of simulations = 30

Fig. 5.27 Comparison of estimates of $E[\hat{P}]$ using simulations and from theory for uniform product kernel. Smoothing space: $u = (N_1)60$, $v=\log(CSRN)$.

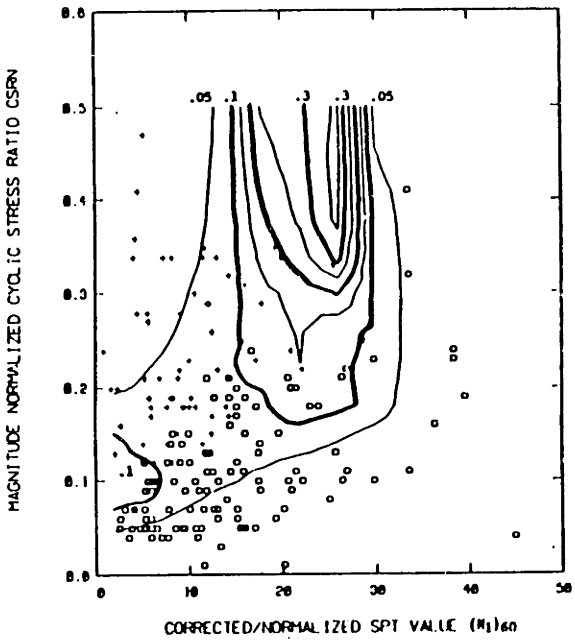


(a) No. of simulations = 30

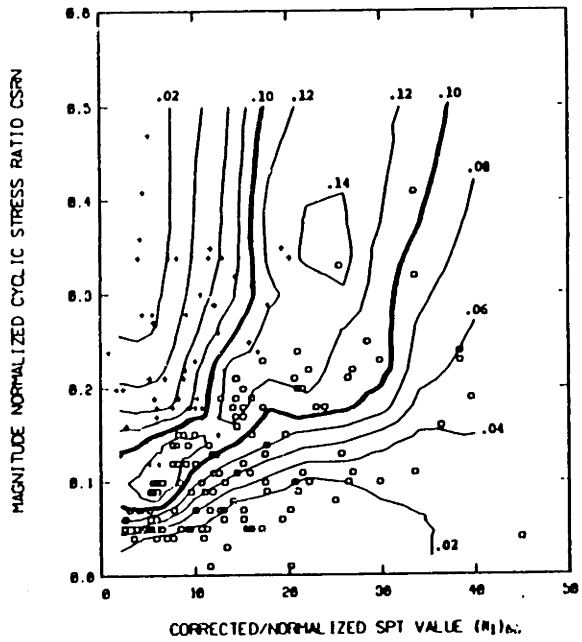
(b) No. of simulations = 60

(c) Theory (No. of simulations → ∞)

Fig. 5.28 Comparison of estimates of standard deviation $SD[\hat{P}]$ using simulations and from theory for uniform product (UP) kernel. Smoothing space: $u = (N_1)60$, v = CSRN.

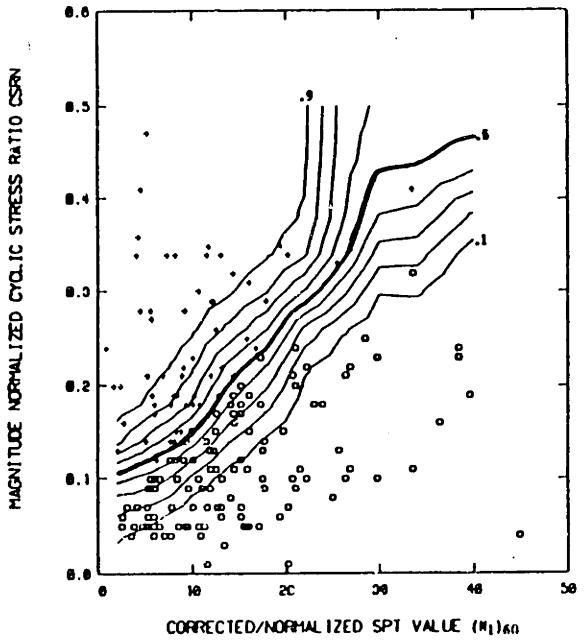


(a) $\widehat{SD}[\hat{P}]$ for uniform product (UP) kernel

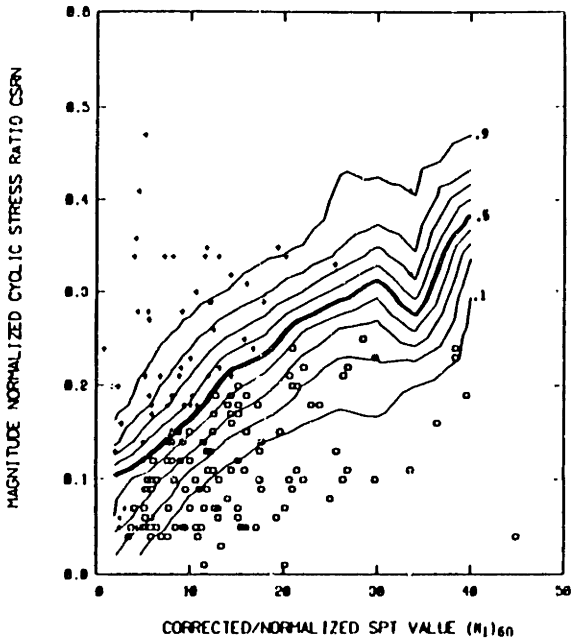


(b) $\widehat{SD}[\hat{P}]$ for variable product (VP) kernel

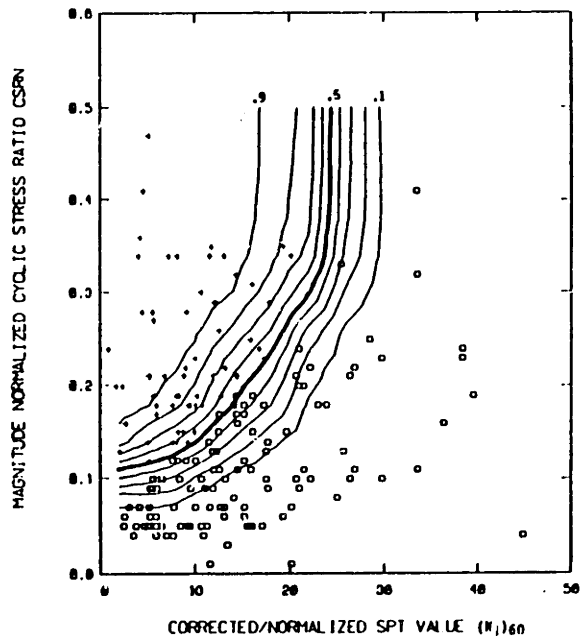
Fig. 5.29 Estimates of the standard deviation of $\hat{P}(Y=1)$ based on actual data (i.e. not assuming underlying P). Contours show equal values of $SD[\hat{P}]$.



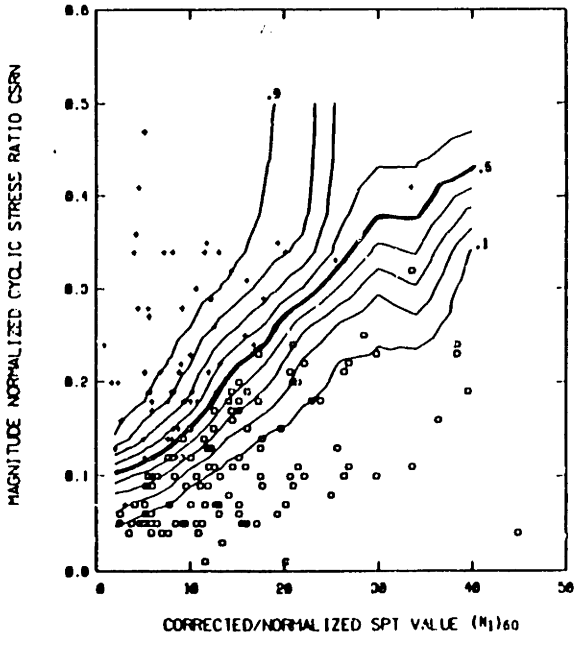
(a) $(N_1)_{60}$ - CSRN logistic model



(b) $\log[(N_1)_{60}]$ - CSRN logistic model



(c) $(N_1)_{60}$ - log(CSRN) logistic model



(d) $\log[(N_1)_{60}]$ - log(CSRN) logistic model

Fig. 5.30 Comparisen $\hat{P}(Y=1)$ obtained assuming various underlying logistic models for P. Kernel used was the optimum uniform product kernel found for estimating $\hat{P}$ using the actual data. Smoothing space: $u = (N_1)_{60}$, $v = \log(CSRN)$. (Refer to Fig. 5.11 for actual P models.)

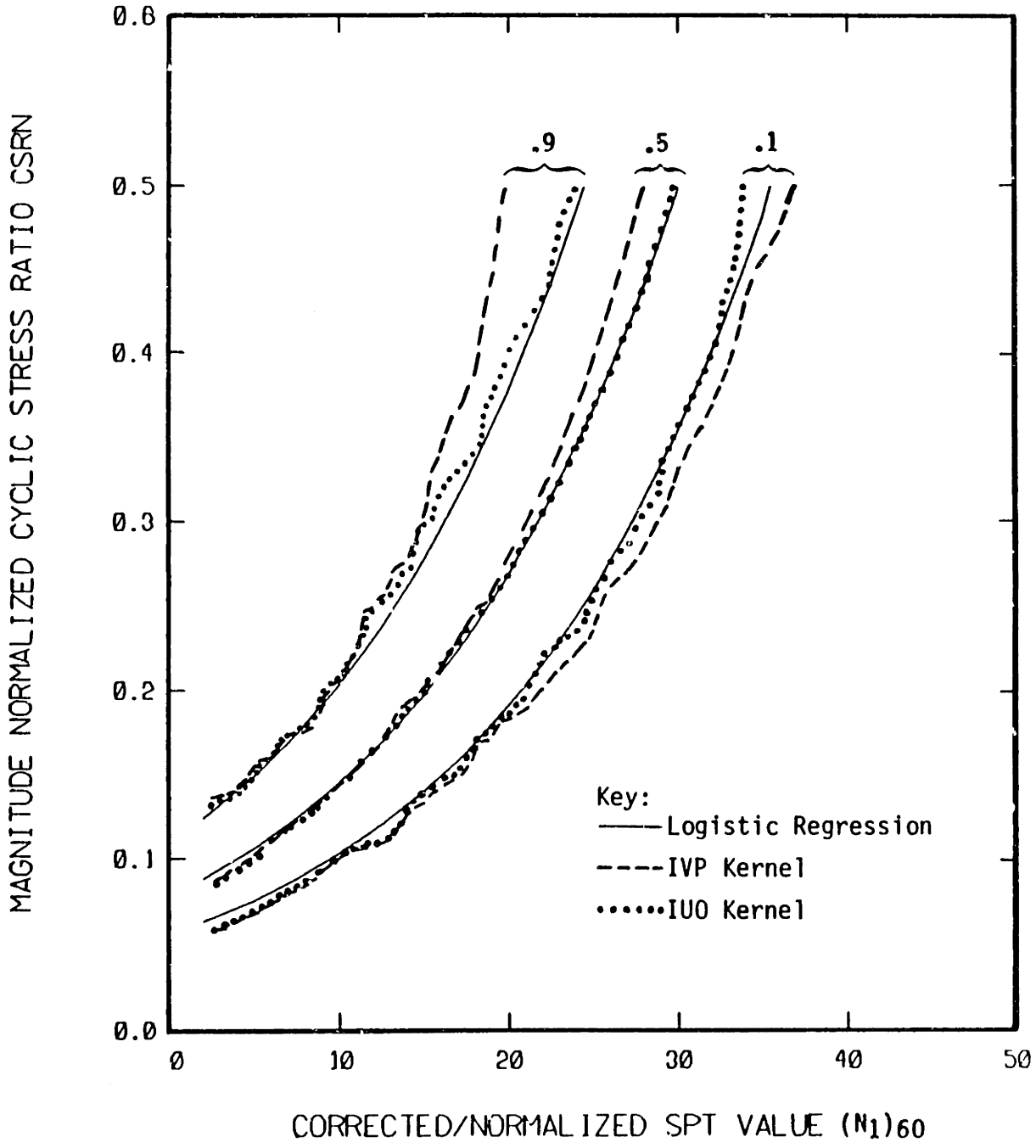


Fig. 5.31 Superimposed comparison of the P = 0.1, 0.5 and 0.9 equi-probability contour lines obtained by logistic regression and by nonparametric kernel estimation. (Kernel smoothing space: u = $(N_1)_{60}$, v = log(CSRN)).

CHAPTER 6

ANALYSIS OF ERRORS AND BIASES IN THE DATA BASE

6.1 Introduction

In Chapter 4, the liquefaction data catalog was analyzed using standard logistic regression techniques to identify the relevant explanatory variables and to formulate parametric models of liquefaction behavior. In Chapter 5, a non-parametric approach was employed which gave further support to the parametric logistic models. However, in both Chapters 4 and 5, the data were analyzed with the following implicit assumptions:

1) that the data are error-free, i.e., the explanatory variables (e.g., CSRN and $(N_1)_{60}$) are measured exactly,
2) that the responses (liquefaction or no liquefaction) for the various cases are independent random variables, and
3) that the documentation and reporting of case studies in the literature is not influenced by whether or not a site liquefied, so that the proportion of liquefaction to non-liquefaction observations reflects the actual proportion in nature.

As discussed in Chapter 3, the above assumptions are not entirely justified in light of the methodology used in assembling the data catalog. The purpose of the analyses presented in this chapter is to demonstrate the effects of violation of these assumptions on the results of logistic regression.

Various modifications of the standard logistic regression procedure

are used to re-analyze the data. In most cases, the underlying method is that of weighted logistic regression, whereby a different weight is assigned to each data point to reflect the amount of "information" attributed that data point. For example, data points or case studies that are considered to be reliable observations would be weighted more than those considered to be less reliable.

A problem with the use of weights is that it obscures the interpretation of some goodness-of-fit indicators in logistic regression. Another limitation is the judgemental nature of weight assignment. For example, it may be difficult to determine whether an "unreliable" data point should be weighted 1/2 or 1/10 as much as a "reliable" data point. Thus the results of this chapter should not be viewed as definitive quantitative corrections to the standard logistic results, but rather as indications of qualitative trends caused by the data imperfections. These trends are useful in identifying potential biases (e.g. towards conservatism or unconservatism) of the logistic models.

6.2 Weighted Logistic Regression

Weighted logistic regression is a simple extension of the standard logistic regression procedure described in Chapter 4. In logistic regression, the objective is to estimate the probability P that the binary response Y equals 1 through an expression of the type:

$$P = 1/[1 + \exp -\{\underline{x}^T\underline{\beta}\}] \qquad (6.1)$$

where $\underline{x}$ is a vector of explanatory variables and $\underline{\beta}$ is a vector of coefficients to be obtained by fitting Eqn. (6.1) to the data. The method of maximum likelihood is used to estimate $\underline{\beta}$; see Chapter 4. In

standard logistic regression, the log-likelihood is:

$$L = \sum_{i-1}^{n} [Y_i \ln P_i + (1-Y_i) \ln(1-P_i)] \tag{6.2}$$

where Y_i denotes the ith observation (Y_i=1 or Y_i=0), and P_i is obtained from Eqn. 6.1 for $\underline{x}=\underline{x}_i$. In weighted logistic regression, L is rewritten as:

$$L_w = \sum_{i=1}^{n} w_i [Y_i \ln P_i + (1-Y_i) \ln(1-P_i)] \tag{6.3}$$

i.e. the contribution of the ith data point to the log-likelihood is multiplied by a weight w_i. In the extreme case when w_i=0, the ith data point is excluded from analysis. Assigning w_i=1 for all i=1,...n results in the standard logistic regression.

The Newton-Raphson algorithm described in Sec. 4.3 is easily adapted to maximization of the weighted log-likelihood. It can be shown that the first and second derivatives of the weighted log-likelihood required for the algorithm are:

$$\frac{\partial L_w}{\partial \beta_k} = \sum_{i=1}^{n} w_i \, x_{ik}(Y_i-P_i) \tag{6.4}$$

$$\frac{\partial^2 L_w}{\partial \beta_i \partial \beta_k} = \sum_{i=1}^{n} -w_i \, x_{ij} \, x_{ik}(1-P_i)P_i \tag{6.5}$$

where x_{ik} denotes the kth element of the data vector $\underline{x}_i$.

Goodness-of-fit statistics such as the likelihood ratios (LR_0 and LR_1) and the modified likelihood ratio index (MLRI or $\overline{\rho}^2$) can be calculated based on the weighted likelihood function. However, it is not clear how the weighted likelihood statistics should be interpreted,

particularly for purposes of hypothesis testing. This is due to the fact that the distribution of the test statistics are not known when weights are used. As an example, the number of degrees of freedom for the T-statistic of the parameter estimates are clearly affected by the weighting, but it is not clear how to calculate the equivalent degrees of freedom for the weighted regression. Direct comparison of weighted statistics with those from a standard logistic regression are also inappropriate. As a consequence, weighted regression should not be used in stepwise procedures that employ hypothesis testing to identify the best set of explanatory variables. In this chapter, weighted logistic regression is applied to those models previously identified in Chapter 4 as statistically most significant.

6.3 Variable Data Quality and Dependence

The problems of variable data quality and probabilistic dependence among the data were treated by using weighted logistic regression with rationally derived weighting schemes. In formulating these weighting schemes, one must determine:

(a) which data or groups of data are less reliable or are correlated, and

(b) how much weight should be given to such data.

This is not a simple task because of lack of information regarding specific case studies and because of lack of knowledge regarding the physical process of liquefaction, e.g. how far does excess pore pressure propagate in a soil layer so that liquefaction at one location is actually triggered by liquefaction at another location? Since the

assignment of weights to the data inherently involves judgment, not one but several weighting schemes were used.

The rationale followed in devising the weights to reflect variable data quality and lack of independence was based on the notion of information content of the data. Consider, for example, the use of data from two soil borings. If the two borings are relatively far apart, then the SPT resistances measured in the borings can be considered independent. On the other hand, if the borings are extremely close to each other and the profiles of SPT resistances are nearly identical, then the "information" contained in one of the borings is clearly redundant. It would thus be inappropriate to use data from the two borings as independent pieces of information. One should either exclude one of the borings from the analysis or, alternatively, assign each data point a weight of 1/2 if both data are retained in the analysis. Similarly, one can argue that a less reliable data point should be given a smaller weight than a more reliable data point.

Two criteria for identifying relatively less reliable data have been used. One criterion considers data prior to 1964 as less reliable. This is the year of the Niigata and Alaskan earthquakes, which brought the liquefaction problem to wide attention of the geotechnical engineering profession. The second criterion for segregating the data into groups of different reliability uses a documentation code (DOC) tabulated in the Catalog of Appendix A. This code reflects the amount of information on soil boring data available to the writer; for example, primary source references in Japanese relevant to several cases originally compiled by Toksimatsu and Yoshimi (1983) were not available. However, because of

the somewhat arbitrary character of this criterion, the fact that a particular case study is classified as poorly documented does not necessarily mean that it is intrinsically unreliable.

There are several possible sources of correlation among the case studies. The most evident is when two or more borings are in close geographic proximity to each other. However, the distance between borings are seldom reported and a judgement has to be made as to the degree of correlation between the borings or, conversely, the equivalent number of independent cases represented by a group of borings. For purposes of analyses, upper and lower bounds were estimated for the number of geographically correlated borings that experienced the same earthquake. The lower bound was usually the number of borings known to have been drilled in the same soil deposit, whereas the upper bound was estimated more judgementally, considering the lack of documented information on the exact boring locations. Correlations arising from using the same boring serially in a succession of earthquakes as different case studies were also noted in the data catalog. A summary of the tabulated codes used to formulate weighting schemes is presented in Table 6.1. More detailed explanations for some of the codes may be found in Appendix A.

The various weighting schemes and the results of logistic regression are shown in Figures 6.1 and 6.2 for the Seed-Idriss model with the clean sand data subset. The overall conclusion is that the logistic regression parameters do not vary significantly as a function of the weighting scheme. A few multiplicative combinations of the weights were also tried with similar results. The range of contours of equal probability of

liquefaction which result from the various weighting schemes are shown in Fig. 6.3. Weighting schemes identical to those shown in Figures 6.1 and 6.2 were applied to the silty sand data and to the Davis-Berrill (Epicentral) and Modified Davis-Berrill (Hypocentral) models. Small effects were observed also for these regressions, as illustrated in Figures 6.4, 6.5, and 6.6. Thus, the conclusion drawn from the present analyses is that the non-uniform reliability and the possible correlations among the data do not significantly affect the logistic regression models.

6.4 Biases Due to the Sampling Plan

6.4.1 Ideal and Imperfect Sampling

There are several potential biases in the liquefaction data base due to the methodology of data collection. This methodology, which determines where and how many borings are drilled, is referred to as the "sampling plan". A schematic of how an "ideal" sampling plan can be constructed is shown in Fig. 6.7. First, a geographic area of interest is defined and regions of liquefaction-susceptible soils are identified and mapped. Then borings are drilled on a regular grid in each liquefaction susceptible region, so that the number of borings per unit area is approximately constant. The spacing of the borings should be such that correlations between the borings are small. Preferably, the borings should be drilled prior to the occurrence of an earthquake to avoid the potential problem of changes occurring in the SPT N-value due to earthquake shaking. After the earthquake, manifestations of the occurrence of liquefaction would be noted and mapped, e.g. resulting in the pattern shown in Fig. 6.7. Repeating this sampling plan in many

geographic areas and for several earthquakes would provide an ideal data base for logistic regression.

There are two important features of a data base obtained using the ideal sampling plan. These features are a result of the fact that the locations of the borings do not depend on whether the site liquefied or not. First, the proportion of case studies (boring locations) where liquefaction occurred would be roughly the true proportion of liquefied areas (shaded sub-regions) to unliquefied areas (unshaded sub-regions). Secondly, the proportion of liquefied to non-liquefied sites would tend to decrease as one moves further away from the earthquake source zone. The significance of this second feature is that the number of anomalous sites -- sites close to the earthquake epicenter that have not liquefied or sites far away that have liquefied -- would not be statistically out of proportion in contributing "information" to the data base as a whole. Anomalous data of this nature can have a major effect on fitted M&R logistic models.

The sampling plans currently used to obtain liquefaction case study data are far from ideal. For obvious reasons one often waits for an earthquake to occur, and then makes a surface reconnaissance of the affected geographic area. Borings are then performed in the liquefied zones and non-liquefied zones, but not in proportion to the actual frequency of liquefaction or non-liquefaction in nature. This sampling scheme is referred to here as "response-based" sampling, i.e. the locations of borings are decided based on prior knowledge of the response (liquefaction or no liquefaction). Equivalent terms such as "choice based" or "retrospective" sampling are used in other applications (Manski

and Lerman, 1977; Prentice and Pyke, 1979).

Areas with susceptible soils that have not liquefied are more difficult to identify, because it is not as easy (from a surface reconnaissance) to distinguish between these areas and those where the soil is not susceptible to liquefaction. Thus, borings tend to be concentrated in liquefied areas, and the fewer borings performed in non-liquefied areas are frequently near liquefied sites. The fact that there may be large unrecognized areas of non-liquefaction in potentially susceptible soils is implicitly ignored. Or, if these areas are identified on the basis of prior knowledge of geologic conditions, relatively fewer borings are allocated to those areas because of the higher interest in studying failure events. This practice causes over-representation of liquefaction cases, especially for small magnitude earthquakes and at large distances from the epicenter.

The ideal sampling plan is clearly impractical and has been presented for illustrative purposes only. A more realistic proposal from which an unbiased logistic regression model can be obtained is recommended below:

1) After an earthquake has occurred, the areas affected by the earthquake should be mapped, noting both liquefied and non-liquefied areas.
2) In the area of interest (e.g., river basin, county, town), estimates should be made of the proportion of liquefied to non-liquefied sites as a function of distance to the earthquake source.
3) In deciding where to allocate borings, one should proceed as

usual.

The rationale for the above sampling plan is that acquiring the data in such a fashion allows the quantification biases in the sample. These biases can then be taken into account by using the correction procedures described in the following sections. Deviations from the above sampling plan are allowable, as long as the sampling strategy is properly documented. For example, in cases where extensive mapping is not possible, the investigator should make a conscious note of how he decides to allocate borings to the liquefied and non-liquefied areas that are known to exist.

6.4.2 Correction for Response-Based Sampling

Manski and Lerman (1977) have studied the effects of response-based sampling on multinomial choice models. For the case of binary logistic regression, their result is that a correction to the logistic model should be made by maximizing the weighted log-likelihood in Eqn. 6.3 with the weights w_i defined as:

$$w_i = \begin{cases} Q_P/Q_S & \text{if } Y=1 \\ (1-Q_P)/(1-Q_S) & \text{if } Y=0 \end{cases} \qquad (6.5)$$

where Q_P is the true (population) proportion of occurrences of Y=1 (liquefaction) and Q_S is the corresponding sample proportion. To illustrate intuitively the logic of such a weighting scheme, suppose that, in nature, liquefaction occurs in 25% of all cases, i.e. $Q_P = 1/4$, but that, in a response-based sampling plan, half of the cases are of liquefaction, i.e. $Q_S = 1/2$. Then the weight which is assigned to

liquefaction data points is (1/4)/(1/2) = 1/2 and the weight assigned to non-liquefaction data points is (1 - 1/4)/(1 - 1/2) = 3/2. The contribution of each observed case of non-liquefaction to the log-likelihood function is thus weighted three times more than each case of liquefaction. This reflects the true percentage in nature of 75% liquefaction and 25% non-liquefaction cases.

Manski and Lerman (1977) also present an asymptotic result (which they credit to personal communication with D. McFadden), showing that the maximum likelihood parameters of the proposed weighted logistic regression can be obtained from standard logistic regression by a simple correction to the constant term β_o. For binary regression, the corrected constant term β_o^* can be writted as:

$$\beta_o^* = \beta_o - \ln \left[\frac{Q_S/Q_P}{(1-Q_S)/(1-Q_P)}\right] \qquad (6.6)$$

The correction of the β_o term of the Seed-Idriss model for clean sand is shown in Fig. 6.8 as a function of Q_P. For reference, the best-fit model for the case of no correction is:

$$\text{logit}(P) = 16.447 + 6.4603 \ln(CSRN) - 0.3976 (N_1)_{60} \qquad (6.7)$$

where the constant term is $\beta_o = 16.447$. Out of 182 cases involving clean sand, 68 cases of liquefaction are recorded making the sampled proportion $Q_S = 0.37$. From Fig. 6.8, we see that for assumed values of $Q_P < Q_S$, β_o^* is less than β_o, while for $Q_P > Q_S$, β_o^* is greater than β_o. Note that for $Q_S = Q_P$, β_o^* equals β_o. The fact that only the constant term β_o is involved in the correction implies that a parallel shift of the

equi-probability contour lines occurs for the corrected model. This is illustrated in Fig. 6.9, where it is assumed that Q_P = 0.1. The upward shift of the contour lines indicate that ignoring the bias from sampling produces conservative estimates of the probability of liquefaction. This, of course, is for the case when Q_P is smaller than Q_S, as suggested in Section 6.4.1. Other logistic models corrected for an assumed Q_P = 0.1 are shown in Figures 6.10, 6.11, and 6.12 and indicate a similar conclusion.

A major difficulty in correcting the logistic models for response-based sampling is that we do not know the value of Q_P. For any given earthquake, the actual proportion of liquefied soils will depend on the earthquake magnitude and also on how the area of interest affected by the earthquake is defined. A value of Q_P = 0.1 is considered by the writer to be intuitively reasonable for the present data set but is not based on a systematic inventory. One point of reference comes from the 1976 Guatemala Earthquake where it was noted that only about 6% of a region defined by the Villalobos Delta was affected by liquefaction (Hoose et al., 1978; Seed et al., 1981). In the areas of Niigata affected by the 1964 earthquake, the percentage of liquefied areas is probably greater, while in other instances, the proportion will be considerably less.

The concept embodied by Q_P is ill-defined in the context of the liquefaction problem and contributes to the difficulty of its quantification. A more useful concept is to consider the following quantity extracted from Eqn. 6.6:

$$Q = \frac{Q_S/Q_P}{(1-Q_S)/(1-Q_P)} \equiv \frac{SR_L}{SR_N} \qquad (6.8)$$

The numerator SR_L is equivalent to the ratio of number of sites sampled to the number of sites observed to have liquefied. The denominator SR_N is the same rati for non-liquefied sites. Suppose that, following an earthquake, a detailed reconnaissance and inventory os made of liquefied and non-liquefied sites in susceptible soil areas, and that an investigator is now faced with a decision as to how to allocate a certain number of borings to be drilled in those areas. The quantity Q reflects his inclinations towards drilling in liquefied versus non-liquefied areas. For example, if the investigator plans to drill at every observed liquefaction site ($SR_L = 1.0$), but he only plans to drill at half the non-liquefied sites ($SR_N = 0.5$), the ratio Q would be $(1.0/0.5) = 2$. The value of $Q_P = 0.1$ used for the calculations shown in Figures 6.9 through 6.12 (combined with $Q_S \simeq 0.4$ for the catalog data) corresponds to $Q \simeq 6$ as calculated from Eqn. 6.8. This indicates a belief by the writer that the propensity in the past towards investigating liquefaction occurrences is about six times greater than that for non-liquefaction.

6.4.3. M&R Bias Correction

The definition of the weights in Eqn. 6.5, as defined by Manski and Lerman (1977) accounts for an overall bias in the collection of the data. However, in practice the sampling bias is a function of the magnitude and distance to the earthquake source. The sampling bias manifests itself as an anomaly characterized by an over-representation of liquefaction cases for small magnitude earthquakes and for large distances from the earthquake source. This variation of bias particularly affects M&R logistic models and is investigated by considering a logistic regression

of the form:

$$P = 1/[1 + \exp-(\underline{x}^T\underline{\beta})] \qquad (6.9a)$$

where

$$\underline{x}^T\underline{\beta} = \beta_o + \beta_1 M + \beta_2 \ln(R_{EP}) + \beta_3 \ln(\bar{\sigma}_v) + \beta_4 (N_1)_{60} \qquad (6.9b)$$

This model, which is based on the epicentral distance R_{EP}, was found to be statistically significant in Chapter 4 and represents a generalization of the Davis-Berrill model.

Correcting for the variation of bias as a function of M and R can be done by an extension of the Manski and Lerman (1977) procedure. The purpose of the extension is to derive weights similar to those in Eqn. 6.5, except that the weights would depend on M and R. The steps in the algorithm are:

1. Estimate $Q_S = Q_S(M,R)$ using a non-parametric kernel formulation as developed in Chapter 6:

$$Q_S(M,R) = \frac{\sum_{i=1}^{m} K_i(M,R)\, Y_i}{\sum_{i=1}^{m} K_i(M,R)} \qquad (6.10)$$

   where the summation is over the data points (i=1,...n), and K denotes the kernel function.

2. Fit the logistic model of Eqn. 6.9 to the data and denote it as $P(M,R,N,\sigma)$ where the subscripts for R, N and σ are left out for convenience of notation.

3. Estimate the true proportion $Q_P(M,R)$ as:

$$Q_P(M,R) = \underset{N,\sigma}{E}\,[P(M,R,N,\sigma)] \qquad (6.11a)$$

$$\simeq \frac{1}{n} \sum_{i=1}^{n} P(M,R,N_i,\sigma_i) \qquad (6.11b)$$

4. Calculate the weights w_j using a modification of Eqn. 6.5:

$$w_j = \begin{cases} Q_{Pj}/Q_{Sj} & \text{if } Y=1 \\ (1-Q_{Pj})/(1-Q_{Si}) & \text{if } Y=0 \end{cases} \qquad (6.12)$$

where Q_{Pj} and Q_{Sj} are Q_P and Q_S evaluated at M and R for the jth data point (j=1,...n), i.e.

$$Q_{Sj} = Q_S(M_j,R_j) \qquad (6.12a)$$

and

$$Q_{Pj} = Q_P(M_j,R_j) \qquad (6.12b)$$

5. Re-fit the logistic model $P(M,R,N,\sigma)$ using the weighted likelihood formulation and the weights calculated in Step 4.

6. Go back to Step 3. Repeat until convergence.

The final result is a logistic model of the form of Eqn. 6.9, but with coefficients β_k corrected for the M&R bias.

The rationale and the key steps of the above algorithm involve the estimates of $Q_S(M,R)$ and $Q_P(M,R)$. The idea behind using Eqn. 6.10 (Step 1) to calculate $Q_S(M,R)$ is to estimate the local relative frequency of liquefaction whereas Eqn. 6.11 (Step 3) estimates $Q_P(M,R)$ as the empirical average over N and σ of the logistic probability of liquefaction. These proportions $Q_S(M,R)$ and $Q_P(M,R)$ are then used to calculate weights (Eqn. 6.11) which embody the same concept as those calculated using Eqn. 6.5. The difference here is that the weights defined by Eqn. 6.12 are a function of magnitude and distance, whereas

the weights in Eqn. 6.5 are not.

A clarification of the procedure and a display of the results is shown in Fig. 6.13. Part (a) of the figure shows the estimate of $Q_S(M,R)$ obtained using the optimum cross-validated likelihood kernel. The initial estimate of $Q_P(M,R)$ is shown in Fig. 6.13(b). In the first iteration, the weights for a liquefaction point are calculated as the value of $Q_P(M,R)$ from Fig. 6.13(b) divided by the value of $Q_S(M,R)$ from Fig. 6.13(a). The weight for a non-liquefaction point is the ratio between $(1-Q_P)$ and $(1-Q_S)$ where Q_P and Q_S are also obtained from Figures 6.13(b) and 6.13(a). The values of $Q_S(M,R)$ in Fig. 6.13(a) remain constant through the iterations. However, $Q_P(M,R)$ changes at each iteration and its contour lines at convergence are shown in Fig. 6.13(c). By examining Fig. 6.13, it is evident that any anomalous points tend to be assigned a small weight by the proposed algorithm.

Table 6.2 presents a comparison of the bias corrected model and the standard model. It is clear that the major changes are for the coefficients β_1 and β_2 which correspond to the variables M and $\ln(R_{EP})$. The interpretation of the increased magnitude of the β_1 and β_2 coefficients is that the variables M and R_{EP} are actually better determinants of liquefaction occurrence than is currently reflected by the data base with its biases due to the sampling plan. The change in β_o is simply a shift to accommodate the increased magnitudes of β_1 and β_2. The coefficients β_3 and β_4 show hardly any change.

The implications of these changes in the logistic coefficients are illustrated in Fig. 6.14 for the case of $(N_1)_{60} = 10$ and $\bar{\sigma}_v = 0.7$ kg/cm^2. The equi-probability lines for the corrected model (Fig. 6.14b) are much

closer together than those for the standard model due to the increase in the β_1 and β_2 coefficients. However, the slopes of the lines in the two models are practically equal. This is due to the fact that the ratio of β_1 to β_2 has remained essentially the same. It should be noted that this ratio is also roughly the same as that embodied in the Davis-Berrill epicentral model parameter Λ_{EP}, as illustrated by the results of calculations shown in Fig. 6.14(c).

The bias correction procedure presented in this section depends largely on the kernel estimate of $Q_S(M,R)$. Fig. 6.15 shows how the corrected logistic parameters vary as a function of the kernel bandwidth h for a fixed bandwidth ratio $\gamma = 3.1$. As was noted in Table 6.2, the coefficients β_3 and β_4 and the ratio β_1/β_2 are not affected very much by the bias-correction algorithm. This continues to hold true for various kernel sizes as shown in Fig. 6.15(d), (e) and (f). However, there are meaningful (although not large) effects of h on the β_o, β_1 and β_2. Also, as discussed in Chapter 5, it is known that the fixed-size kernel estimator is itself biased for small sample sizes, and this bias can affect the β_k estimates. Use of a variable size kernel would have led to only slightly different estimates of $Q_S(M,R)$. Use of integrated kernels are inappropriate because these diminish the local nature of the $Q_S(M,R)$ estimate, which is considered essential in implementing the bias correction algorithm.

It should be noted that the procedure described in this section corrects for the relative bias rather than the overall bias which results from disproportionate sampling (Section 6.4.2). Also, if the relative bias can be determined externally without reference to the data, then this

information can be used instead of correction procedure described here.

6.5 Effects of Measurement Errors

It is a commonly accepted fact that there is a large degree of statistical measurement error inherent in the values of the SPT resistance and the earthquake intensity parameters. Some of the sources of these errors have been discussed in Chapter 3. The objective in this section is to show the effects of measurement errors on the fitted logistic models.

6.5.1 Errors-In-Variables (EIV) Formulation

In linear regression analysis, it is usually assumed that the response variable is measured with error, but that the explanatory variables are error-free. To consider regressions where both response and explanatory variables are measured with error is a more difficult problem and is referred to as errors-in-variables (EIV) regression. Various papers published on this subject include a review article by Madansky (1959), and more recent articles by Fuller (1980), Gleser (1981) and Ganse et al. (1983). There are also several recent papers that deal with non-linear models, but the papers by Michalek and Tripathi (1980), Carroll et al (1984), and Stefanski and Carroll (1985) are specifically applicable to logistic regression.

In logistic regression, it is also assumed that the responses are measured exactly, i.e. there is no uncertainty regarding whether or not liquefaction occurred. In fact, there are at least a few instances in the data catalog where the interpretation of the actual response is

questionable. Michalek and Tripathi (1980) have addressed this problem in the context of classification. However, only the effects of measurement errors in the explanatory variables are considered here.

In this section, we adopt an approach similar to that presented by Carroll et al. (1984) in which it is assumed that the distribution of measurement error is known a priori. A linear logistic model of the form:

$$\text{logit}(P^*) = (\underline{x}^*)^T\underline{\beta}^* = (\underline{x} - \underline{\varepsilon})^T\,\underline{\beta}^* \tag{6.14a}$$

is assumed, where $\underline{x} = \underline{x}^* + \underline{\varepsilon}$ is the vector of measured explanatory variables, $\underline{\varepsilon}$ is the vector of measurement errors in $\underline{x}$, $\underline{x}^*$ is the vector of true explanatory variables, and $\underline{\beta}^*$ is the parameter vector for measurements without error. Specifically,

$$\underline{\beta}^* = \begin{bmatrix} \beta_0^* \\ \beta_1^* \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_m^* \end{bmatrix} \quad \text{and} \quad \underline{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} \tag{6.14b}$$

and $\underline{\varepsilon}$ is defined as:

$$\underline{\varepsilon} = \begin{bmatrix} 0 \\ \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_m \end{bmatrix} \tag{6.14c}$$

in which the ε_j (j=0,1,...m) are assumed to be independent, and where

the zero entry reflects the lack of error associated with the constant term of the logistic model. The likelihood function is written as:

$$\ell = \prod_{i=1}^{n} \ell_i \tag{6.15}$$

where ℓ_i is the likelihood contribution of the ith observation: $\ell_i = P_i^*$ if $Y_i = 1$, and $\ell_i = 1 - P_i^*$ if $Y_i = 0$. Since the observations are assumed to be independent, the expectation of the likelihood can be expressed as:

$$E[\ell] = \prod_{i=1}^{n} E[\ell_i] \tag{6.16}$$

Define T_i as:

$$T_i = \int_{\underline{\varepsilon}} P_i^*(\underline{x}_i - \underline{\varepsilon}) \cdot f_i(\underline{\varepsilon})\, d\underline{\varepsilon} \tag{6.17a}$$

$$= \int_{\underline{\varepsilon}} [1 + \exp\{-(\underline{x}_i - \underline{\varepsilon})^T \underline{\beta}^*\}]^{-1} f_i(\underline{\varepsilon})\, d\underline{\varepsilon} \tag{6.17b}$$

In the above equation $f_i(\underline{\varepsilon})$ denotes the probability distribution of the measurement error and the notation is a shorthand for

$$f_i(\underline{\varepsilon}) = f(\underline{\varepsilon} | \underline{x}_i) \tag{6.18}$$

in which we allow the measurement error to be a function of $\underline{x}_i$. T_i^* is the expected value of P_i^* (given by Eqn. 6.14) for a given error distribution $f_i(\underline{\varepsilon})$ for $Y_i = 1$, i.e. $E[\ell_i] = T_i$. Similarly, it is easy to show that $E[\ell_i] = 1 - T_i$ for $Y_i = 0$. The resulting log of the expected

likelihood can be written as:

$$L_E = \sum_{i=1}^{n} [Y_i \ln T_i + (1 - Y_i) \ln(1 - T_i)] \qquad (6.16)$$

This likelihood function is maximized to obtain the $\underline{\beta}^*$ parameters.

6.5.2 Implementation

The estimation of the logistic regression coefficients using the errors-in-variables (EIV) formulation is illustrated for the Seed-Idriss model:

$$\text{logit}(P) = \beta_o + \beta_1 \ln(\text{CSRN}) + \beta_2 (N_1)_{60} \qquad (6.20)$$

The objective is to derive the coefficients β_o^* , β_1^* , and β_2^* for the for the situation where the errors ε_1 and ε_2 have been removed from the measurements of ln(CSRN) and $(N_1)_{60}$ using the model:

$$\begin{aligned} \text{logit}(P^*) &= \beta_o^* + \beta_1^*[\ln(\text{CSRN}) - \varepsilon_1] + \beta_2^*[(N_1)_{60} - \varepsilon_2] \\ &= \beta_o^* + \beta_1^* \ln(\text{CSRN}^*) + \beta_2^*[(N_1)_{60}^*] \end{aligned} \qquad (6.21)$$

In the analyses presented, it is assumed that the errors ε_1 and ε_2 are independent and normally distributed with zero means. Various values have been assumed for the standard deviation of ε_1 (denoted as σ_1) and the coefficient of variation of $(N_1)_{60}$ (denoted as V_2, i.e. $V_2 = \sigma_2/E[(N_1)_{60}]$). The cyclic stress ratio CSRN is directly proportional to the peak ground acceleration 'a' which is the parameter that dominates the error variance σ_1. Since the distribution of 'a' is considered to be

approximately log-normal about the median value given by a ground motion attenuation relationship (e.g. see Cornell, 1971), it is therefore not unreasonable to assume that the ε_1 has a log-normal distribution. The measurement error in $(N_1)_{60}$ was modelled by a constant coefficient of variation V_2 because this seemed intuitively reasonable.

Maximization of the likelihood function L_E is obtained through the use of an algorithm in the IMSL library (subroutine ZXMIN) which does not require explicit evaluation of the derivatives of L_E. Also, T_i in Eqn. 6.18 is calculated through numerical integration. In our application involving two explanatory variables, with ε_1 and ε_2 independent and normally distributed, the integration in Eqn. 6.18 was performed using a two-dimensional Hermite integration scheme (Abramowitz and Stegun, 1972, p. 924).

6.5.3 Results From Errors-In-Variables Analysis

Table 6.3 presents results of logistic analysis for various assumed levels of error in ln(CSRN) and $(N_1)_{60}$ using the Seed-Idriss Model and clean sand data. The values of $\sigma_1 = 0.5$ and $V_2 = 0.2$ are estimates considered to be somewhat high. The other assumed levels of error are scaled down from these estimates. As the level of assumed measurement errors increases, the absolute values of the $\underline{\beta}^*$ parameters increase. A triplet of calculations for combinations of $\sigma_1 = 0$ or 0.25 and $V_2 = 0$ or 0.1 indicates that the regression coefficients are sensitive mainly to the measurement error in ln[CSRN].

The implications of the increases in $|\beta_k^*|$ is illustrated in Fig. 6.16. As the level of assumed measurement error increases, the contour

lines of equal probability P^* tends to shrink towards the 0.5 contour. If the assumed errors are sufficiently large, the logistic regression contours collapse to a practically deterministic line separating liquefaction and non-liquefation behavior. This observation is expected because the EIV procedure basically removes part of the uncertainty in the logistic models due to measurement errors. A similar though less severe collapse of the equi-probability contour lines is shown in Fig. 6.17 for the silty sand data for which the coefficients β_k^* were given in Table 6.4. Results obtained for the Davis-Berrill M&R model are similar to those for the Seed-Idriss model.

The application of the EIV logistic model to seismic risk analysis involves consideration of two problems. First, in the estimation of the model parameters, there is currently not enough information to accurately specify the level of measurement error that is involved. The second problem arises in understanding the proper context for the application of the EIV models. Consider a future project site where there is a need to evaluate the probability of liquefaction, and one wishes to decide whether or not to use an EIV model corrected for SPT measurement errors. Two situations can be distinguished. One situation involves measurements of N_1 which are "exact" and without statistical error. In this case, it is important to account for errors in the historical data using the procedures described in this section. The other situation involves measurements of N_1 (at the future site) which have approximately the same measurement errors as the historical data. In this second situation, it would be more expedient to neglect measurement errors entirely and use standard logistic models in analysis and application. More generally,

use of the EIV model is necessary only when there are differences in the error structures of the historical data versus the future measurements.

6.6 Chapter Summary and Conclusions

Various errors and biases of the data base used in this study have been considered, and techniques have been developed to account for these features in the context of logistic regression. Definitive quantitative corrections to the logistic models cannot be made due to the lack of data and information regarding the errors and biases. However, several important conclusions can be drawn from the analyses presented in this chapter:

1) Non-uniform data quality does not significantly affect the logistic regression results. This is probably due in part to the preliminary elimination of data points considered to be unreliable (or outliers) during the compilation of the liquefaction catalog (see Chapter 3).
2) Similarly, the lack of statistical independence between the responses (liquefaction or no liquefaction) of some of the case studies also do not significantly affect the results.
3) There is a potentially severe bias in the prediction of liquefaction probability due to the response-based sampling methodology which is currently used to document the case studies. Use of the standard logistic models (not corrected for this effect) leads to conservative estimates of the probability of liquefaction.
4) The bias caused by response-based sampling depends on the

earthquake magnitude and the site-to-source distance. Removing the bias in the context of M & R logistic models indicates a greater sensitivity of the response to magnitude and distance. However, removing the bias does not appear to affect the slopes of the equi-probability lines in the M versus $\log(R_{EP})$ plane. The slope of these lines is consistent with that implied by the Davis-Berrill model.

5) Statistical errors in the measurement of earthquake intensity and soil resistance (i.e. CSRN and $(N_1)_{60}$) also decrease the predictive ability of the logistic models. The lack of precision of CSRN affects the results more than the lack of precision in $(N_1)_{60}$. Measurement errors taken properly into account may, in some instances, lead to a practically deterministic liquefaction criterion. However, such a deterministic criteria could only be used if actual future measurements with little to no error could be made.

It is clear that biases in the data can significantly affect the logistic models for liquefaction analysis. Many of these biases can be mitigated by the proper collection of data. A practical sampling plan is proposed in which mapping of liquefied and non-liquefied areas is done prior to drilling borings to obtain the SPT resistance. More details are provided in Section 6.4.1.

Table 6.1

Summary of Code Variables Used in Formulating Weights for Logistic Regression Analyses

CODE	DESCRIPTION
DOC	Documentation code: 1 = well documented soil data 2 = adequately documented soil data 3 = poorly documented (or unavailable) soil data
NC1	Lower bound estimate of the number of correlated borings (i.e. borings in the same geographic vicinity)
NC2	Upper bound estimate of the number of correlated borings (i.e. borings in the same geographic vicinity)
NE	Number of earthquake shakings for which the boring is used
IE	ith most recent earthquake for which the same boring is used
NL	Number of times a site (boring location) has been recorded as experiencing liquefaction
IL	ith most recent time that a site has been recorded as experiencing liquefaction
NN	Number of times a site has been recorded as not experiencing liquefaction (during an earthquake)
IN	ith most recent time that a site has been recorded as not experiencing liquefaction

Note: Refer to Appendix A for more detailed explanations of the above codes.

Table 6.2

Comparison of Logistic Coefficients
of M&R Bias Corrected and Standard Model

Explanatory Variable	Coefficient Symbol	Coefficients	
		Standard Model	M&R Bias Corrected
constant	β_o	-12.95	-26.23
M	β_1	3.05	5.92
$\ln(R_{EP})$	β_2	-1.75	-3.32
$\ln(\bar{\sigma}_v)$	β_3	-0.96	-0.89
$(N_1)_{60}$	β_4	-0.21	-0.22

Model: $\text{logit}(P) = \beta_o + \beta_1 M + \beta_2 \ln(R_{EP}) + \beta_3 \ln(\bar{\sigma}_v) + \beta_4 (N_1)_{60}$

R_{EP} in units of kilometers

$\bar{\sigma}_v$ in units of kg/cm^2

Table 6.3

Logistic Regression Coefficients as a Function
of Assumed Errors in Variables
Seed-Idriss Model (Clean Sand Data)

σ_1	V_2	β_o^*	β_1^*	β_2^*
0.0	0.0	16.5	6.46	-0.397
0.1	0.05	18.5	7.24	-0.450
0.0	0.1	17.7	6.88	-0.439
0.25	0.0	32.6	12.7	-0.798
0.25	0.1	42.1	16.3	-1.06
0.50	0.2	2420.0	900.0	-66.9

Notes: σ_1 = Standard Deviation of the Error in ln(CSRN)
V_2 = Coefficient of Variation of $(N_1)_{60}$

Model: $\text{logit}(P^*) = \beta_o^* + \beta_1^* \ln(CSRN^*) + \beta_2^*(N_1)_{60}^*$
where $CSRN^*$ and $(N_1)_{60}^*$ are measured without error.

Table 6.4

Logistic Regression Coefficients as a Function of
Assumed Errors in Variables
Seed-Idriss Model (Silty Sand Data)

σ_1	V_2	β_o^*	β_1^*	β_2^*
0.0	0.0	6.48	2.69	-0.182
0.1	0.05	6.59	2.73	-0.185
0.25	0.1	7.23	2.99	-0.204
0.0	0.2	6.60	2.73	-0.189
0.5	0.0	11.1	4.61	-0.312
0.5	0.2	11.8	4.89	-0.337

Notes: σ_1 = Standard Deviation of the Error in ln(CSRN)
V_2 = Coefficient of Variation of $(N_1)_{60}$

Model: $\text{logit}(P^*) = \beta_o^* + \beta_1^* \ln(CSRN^*) + \beta_2^*(N_1)_{60}^*$
where $CSRN^*$ and $(N_1)_{60}^*$ are measured without error.

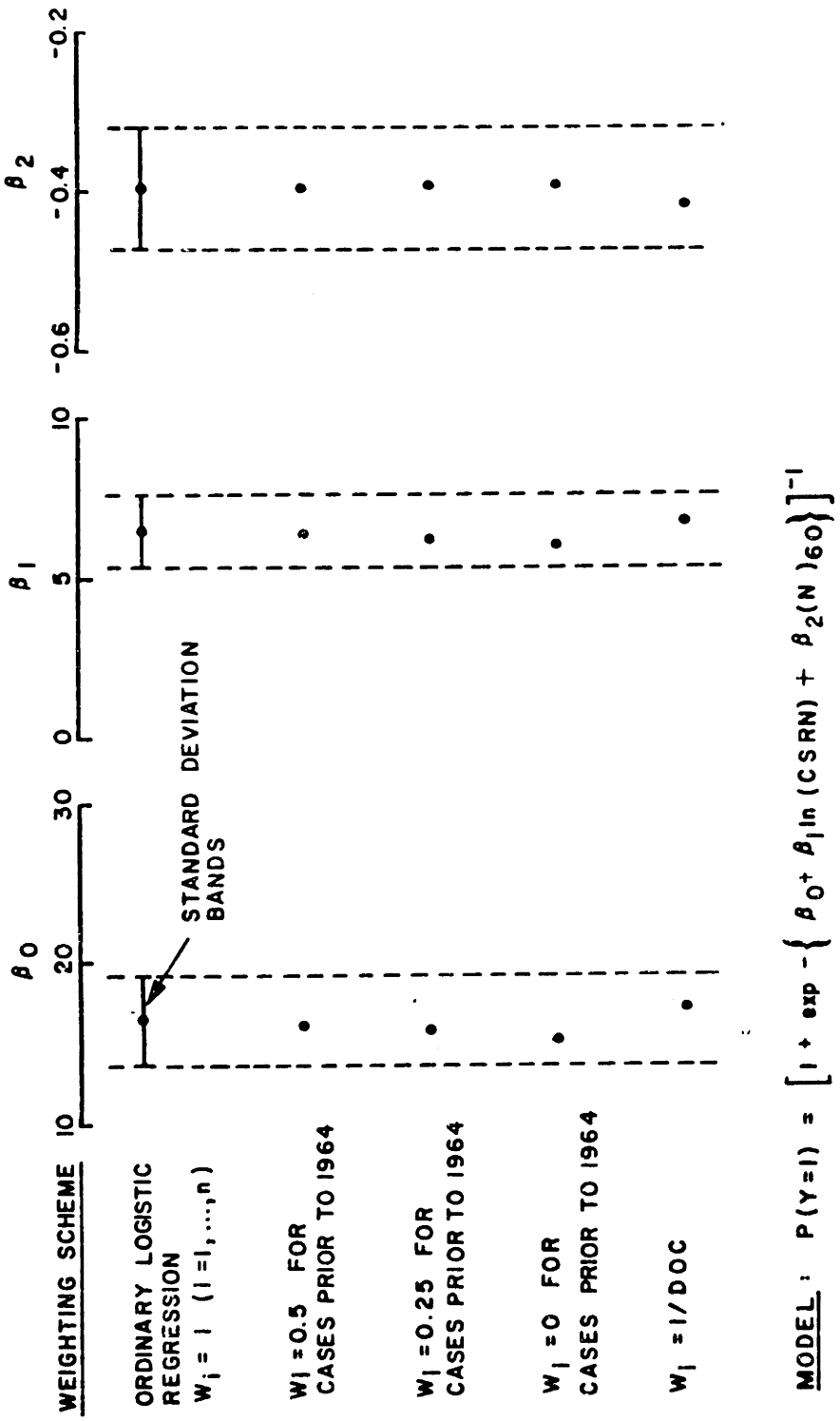


Fig. 6.1 Comparison of logit coefficient estimates obtained using various relative reliability weighting schemes. Seed-Idriss model - clean sand data.

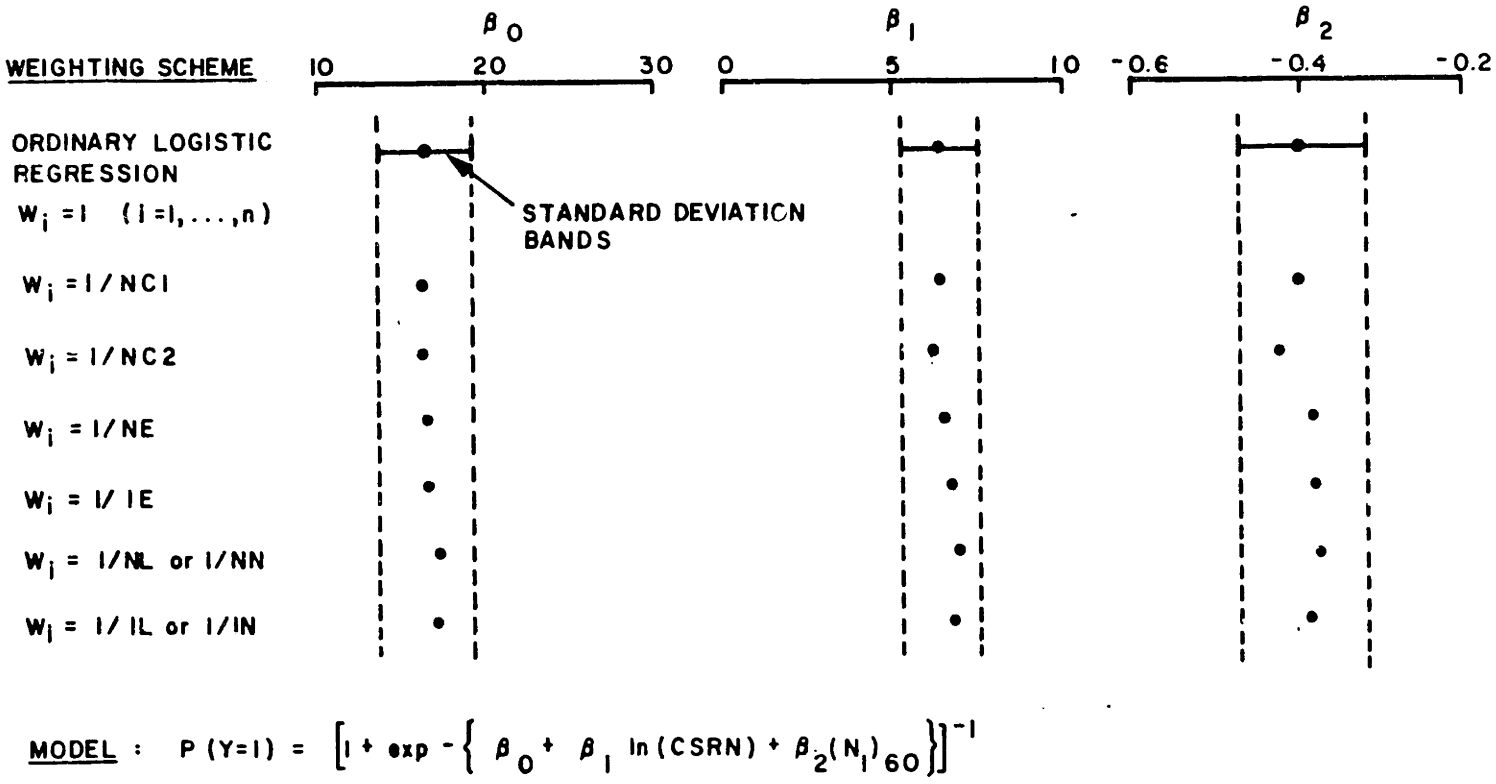


Fig. 6.2 Comparison of logit coefficients obtained using various data independence weighting schemes. Seed-Idriss model - clean sand data. (Refer to Table 6.1 for explanations of weighting formulas.)

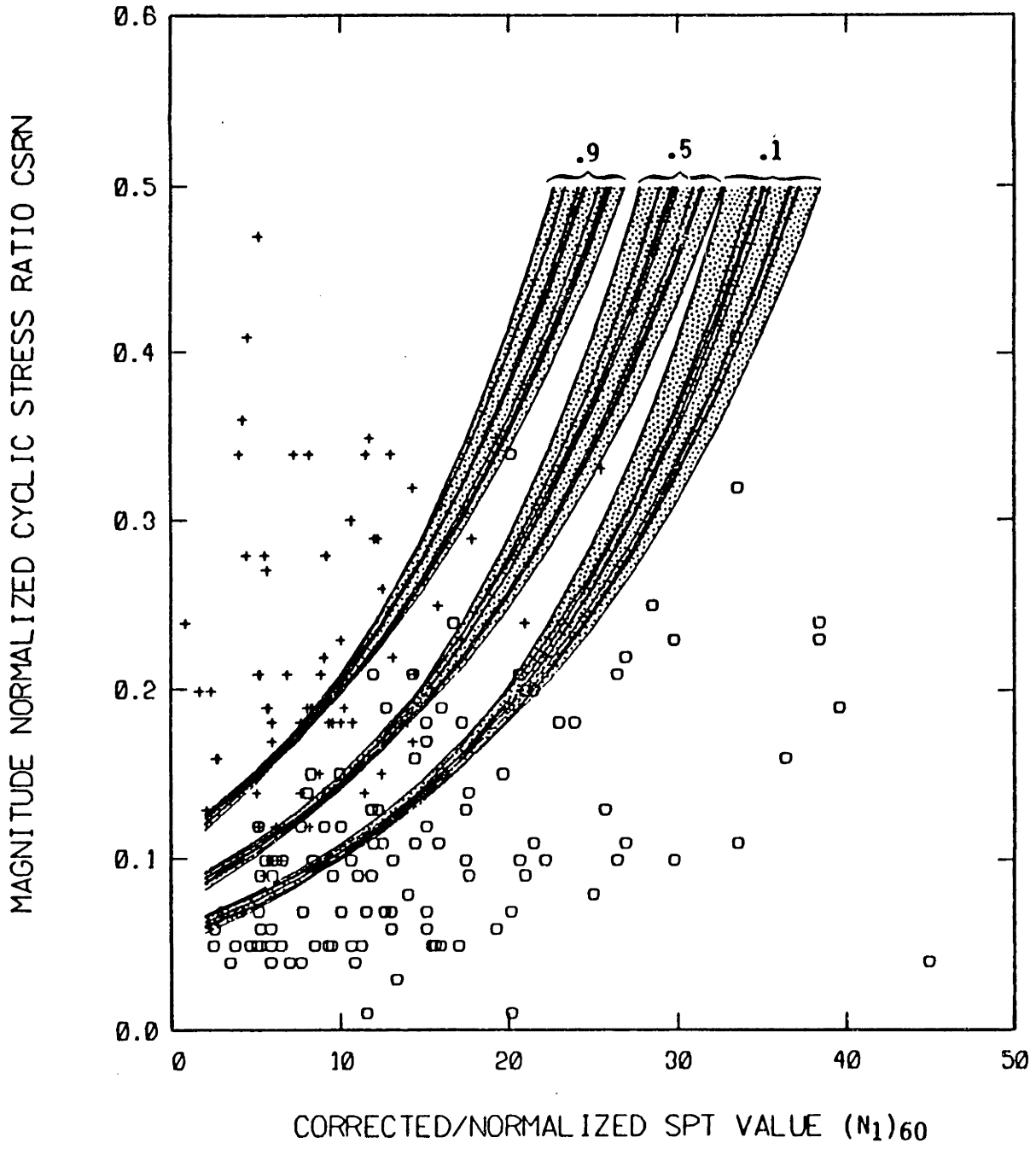


Fig. 6.3 Variation of the 0.1, 0.5, and 0.9 contours of equal probability of liquefaction due to various weighting schemes to account for the variable reliability and non-independence of data. Seed-Idriss model - clean sand data.

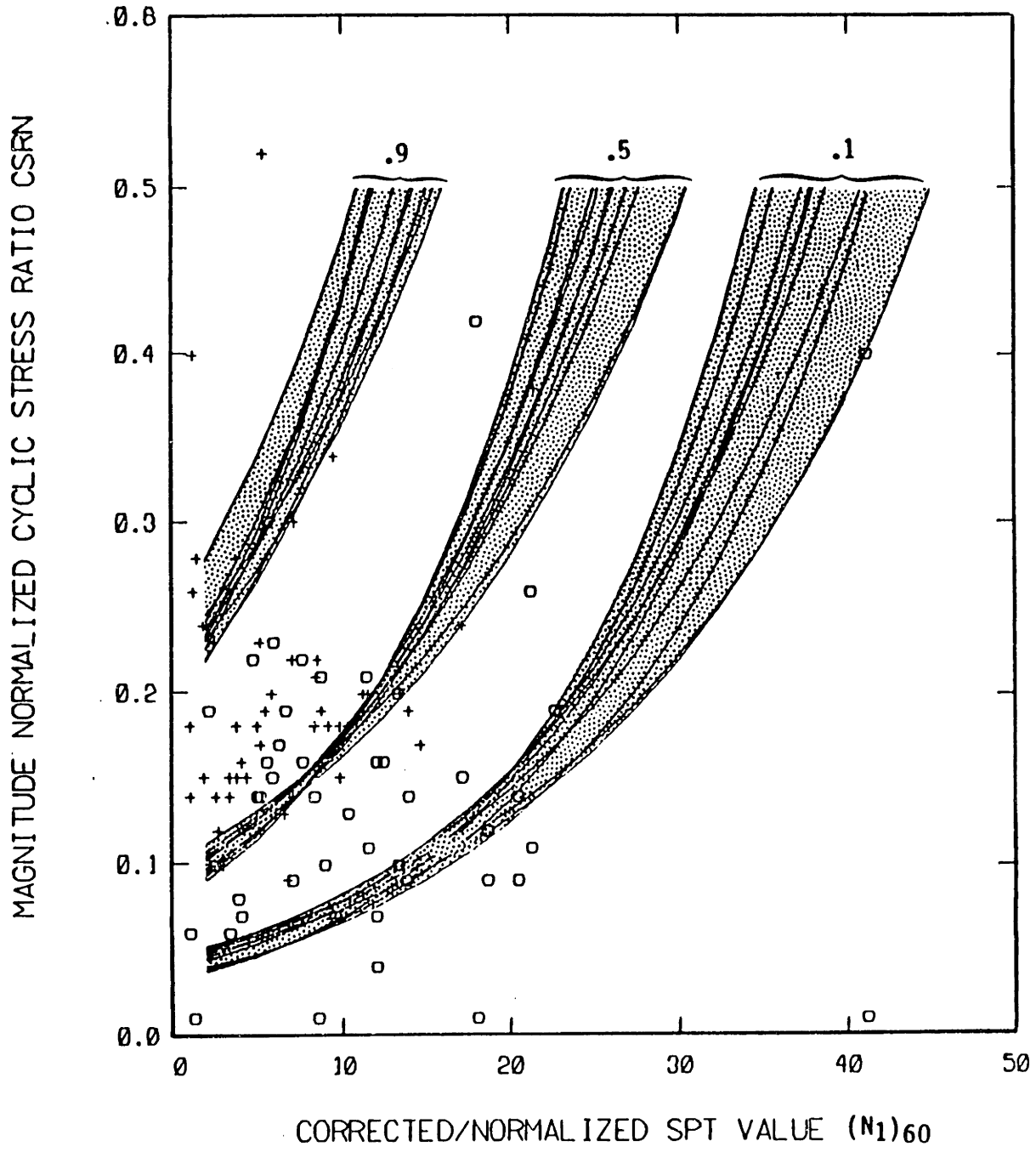


Fig. 6.4 Variation of the 0.1, 0.5, and 0.9 contours of equal probability of liquefaction due to various weighting schemes to account for the variable reliability and non-independence of data. Seed-Idriss model - silty sand data.

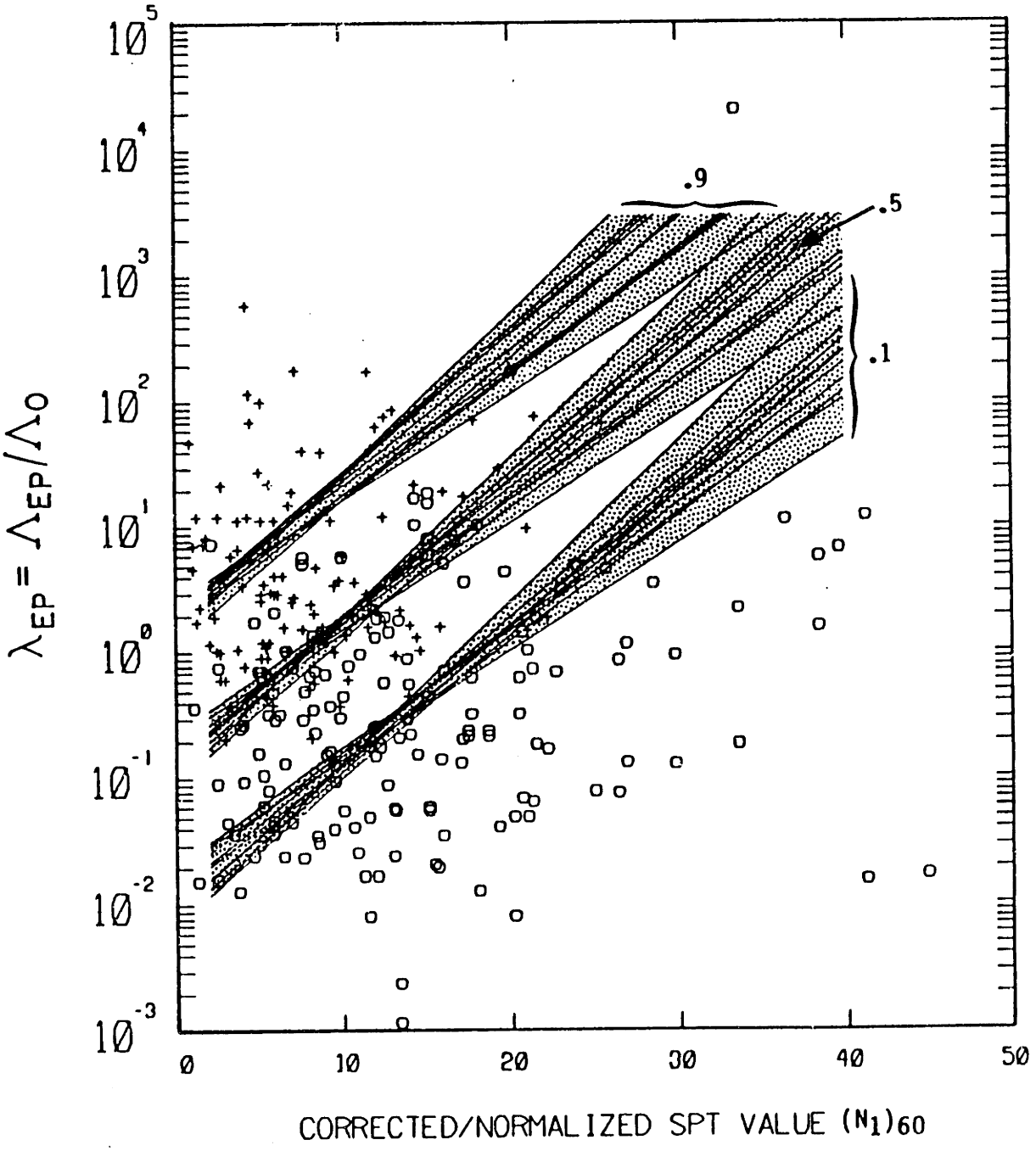


Fig. 6.5 Variation of the 0.1, 0.5, and 0.9 contours of equal probability of liquefaction due to various weighting schemes to account for the variable reliability and non-independence of data. Davis-Perrill epicentral model.

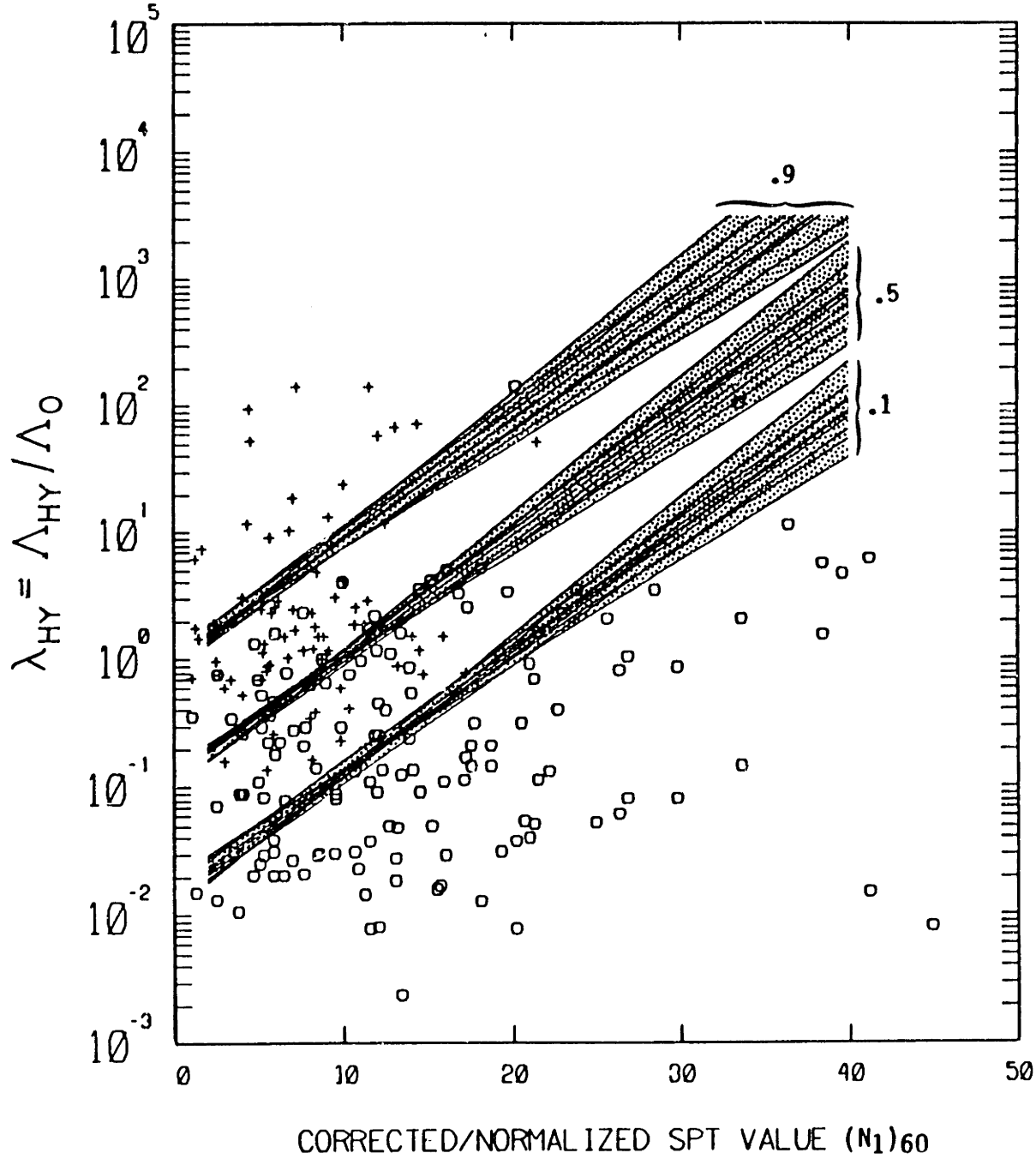


Fig. 6.6 Variation of the 0.1, 0.5, and 0.9 contours of equal probability of liquefaction due to various weighting schemes to account for the variable reliability and non-independence of data. Davis-Berrill hypocentral model.

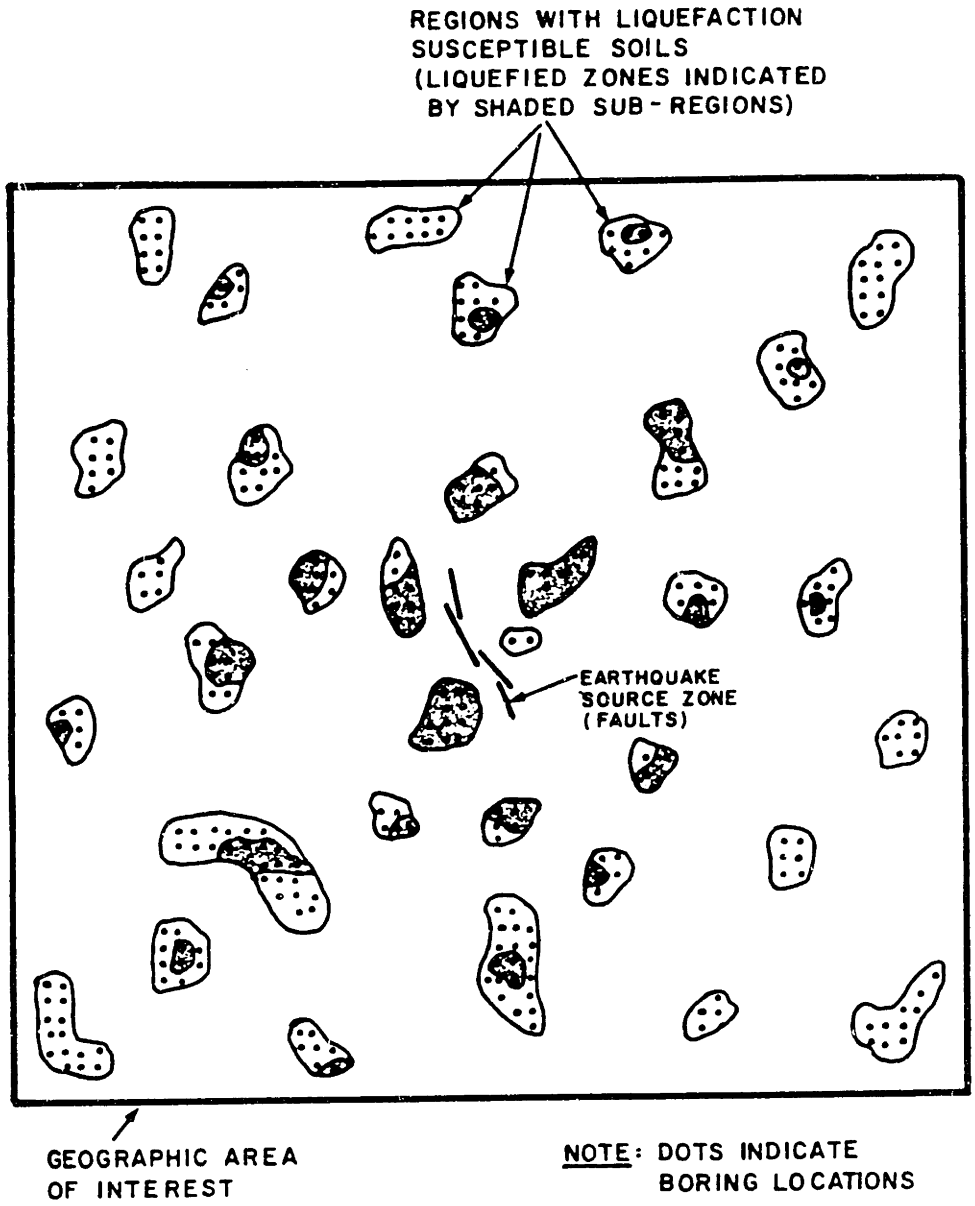


Fig. 6.7 Schematic of an ideal sampling plan and the hypothetical pattern of observed liquefaction after an earthquake.

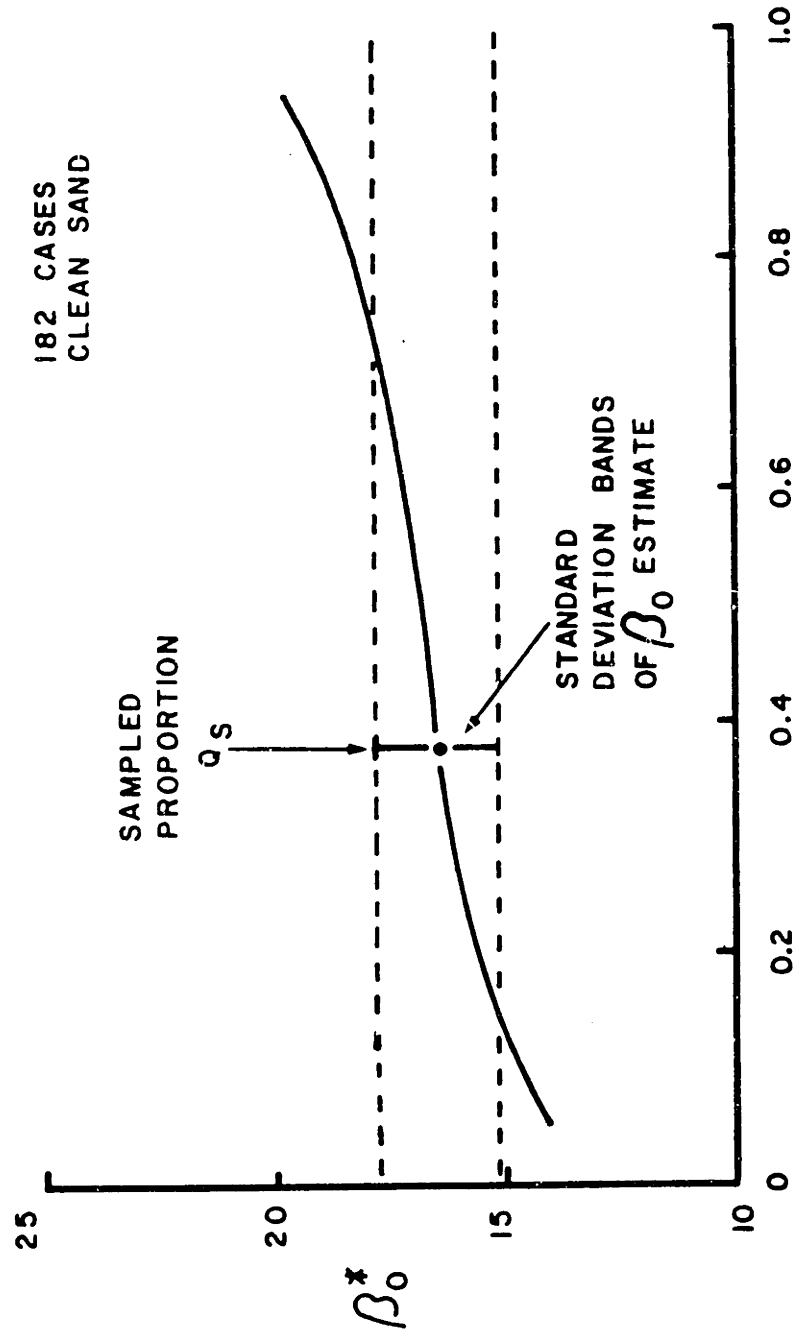


Fig. 6.8 Corrected constant parameter β_0^* as a function of assumed population proportion of liquefied sites Q_P. Seed-Idriss model - clean sand data.

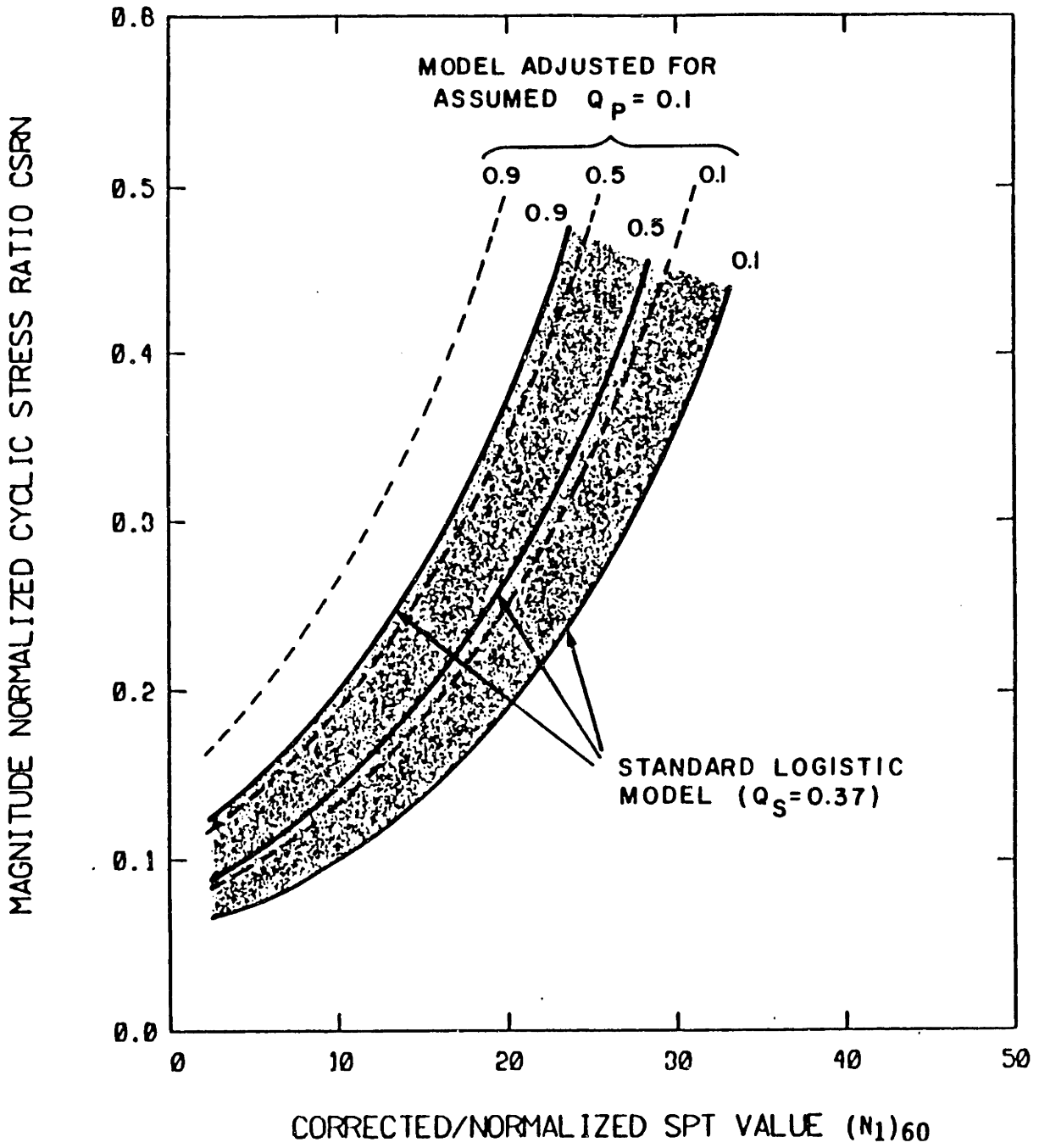


Fig. 6.9 Comparison of 0.1, 0.5, and 0.9 contour lines of equal probability of liquefaction for the standard logistic model and for the model corrected for the effects of choice based sampling (assumed Q_P = 0.1). Seed-Idriss model - clean sand data.

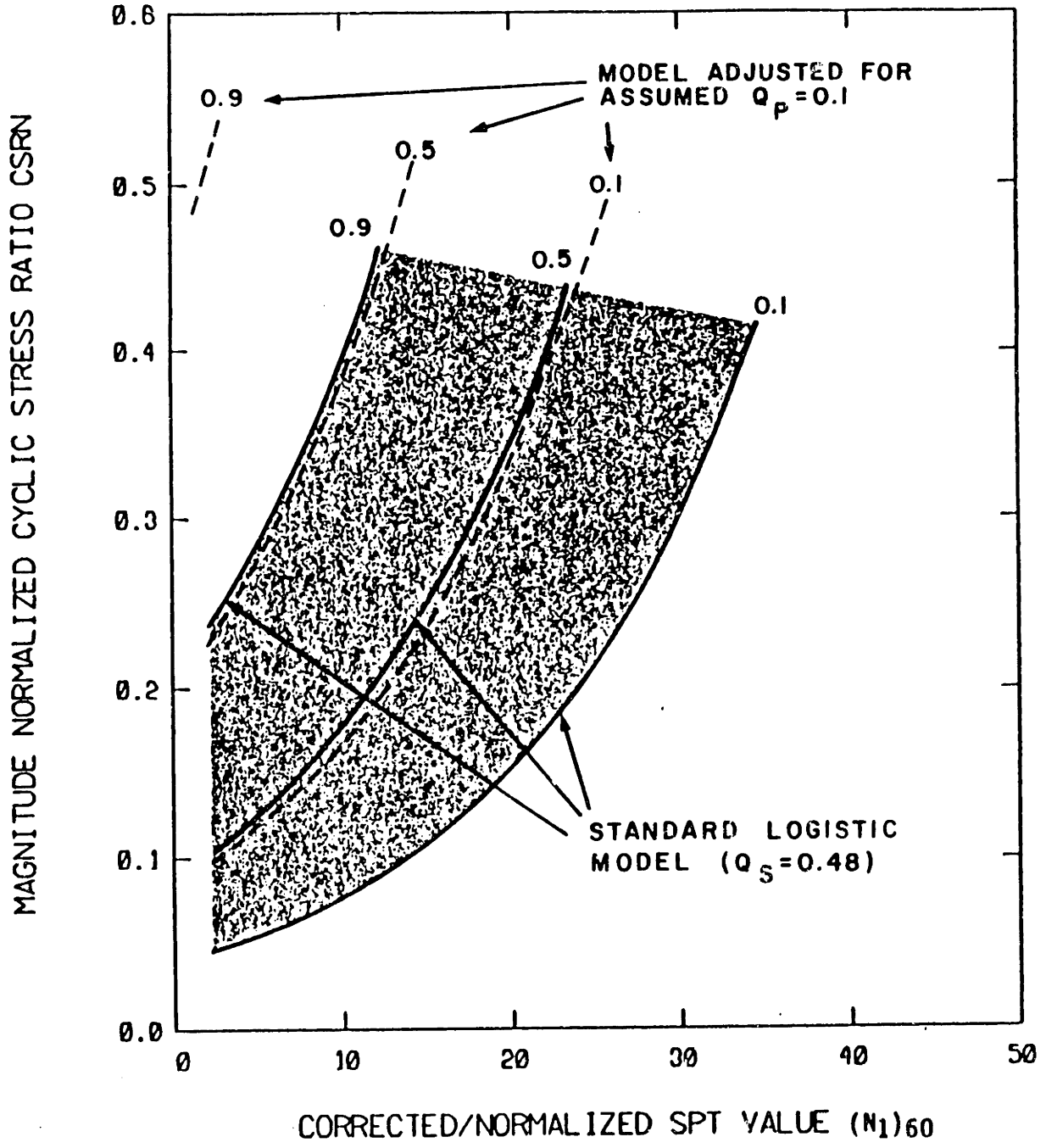


Fig. 6.10 Comparison of 0.1, 0.5, and 0.9 contour lines of equal probability of liquefaction for the standard logistic model and for the model corrected for the effects of choice based sampling (assumed Q_p = 0.1). Seed-Idriss model - silty sand data.

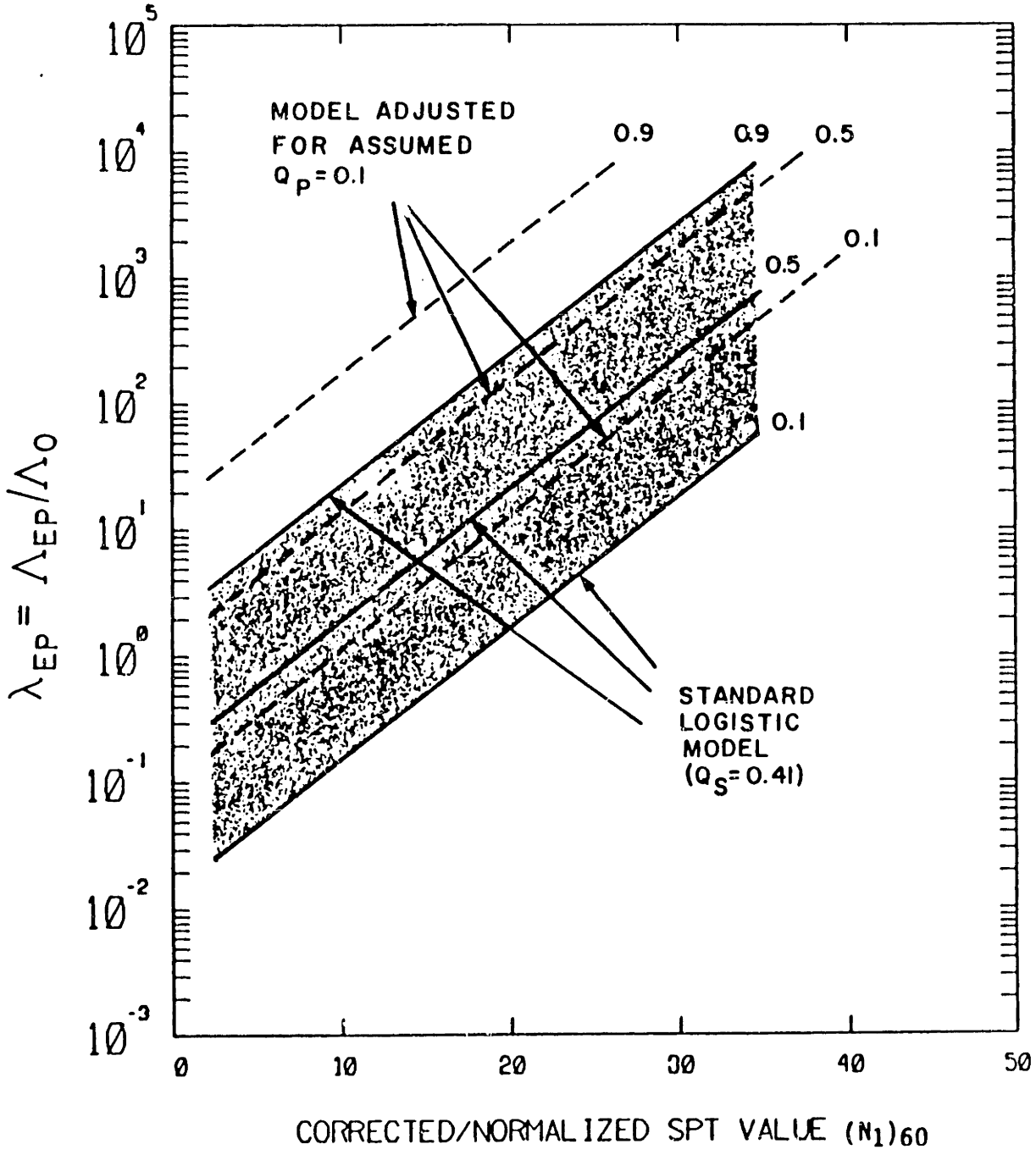


Fig. 6.11 Comparison of 0.1, 0.5, and 0.9 contour lines of equal probability of liquefaction for the standard logistic model and for the model corrected for the effects of choice based sampling (assumed $Q_P = 0.1$). Davis-Berrill epicentral model.

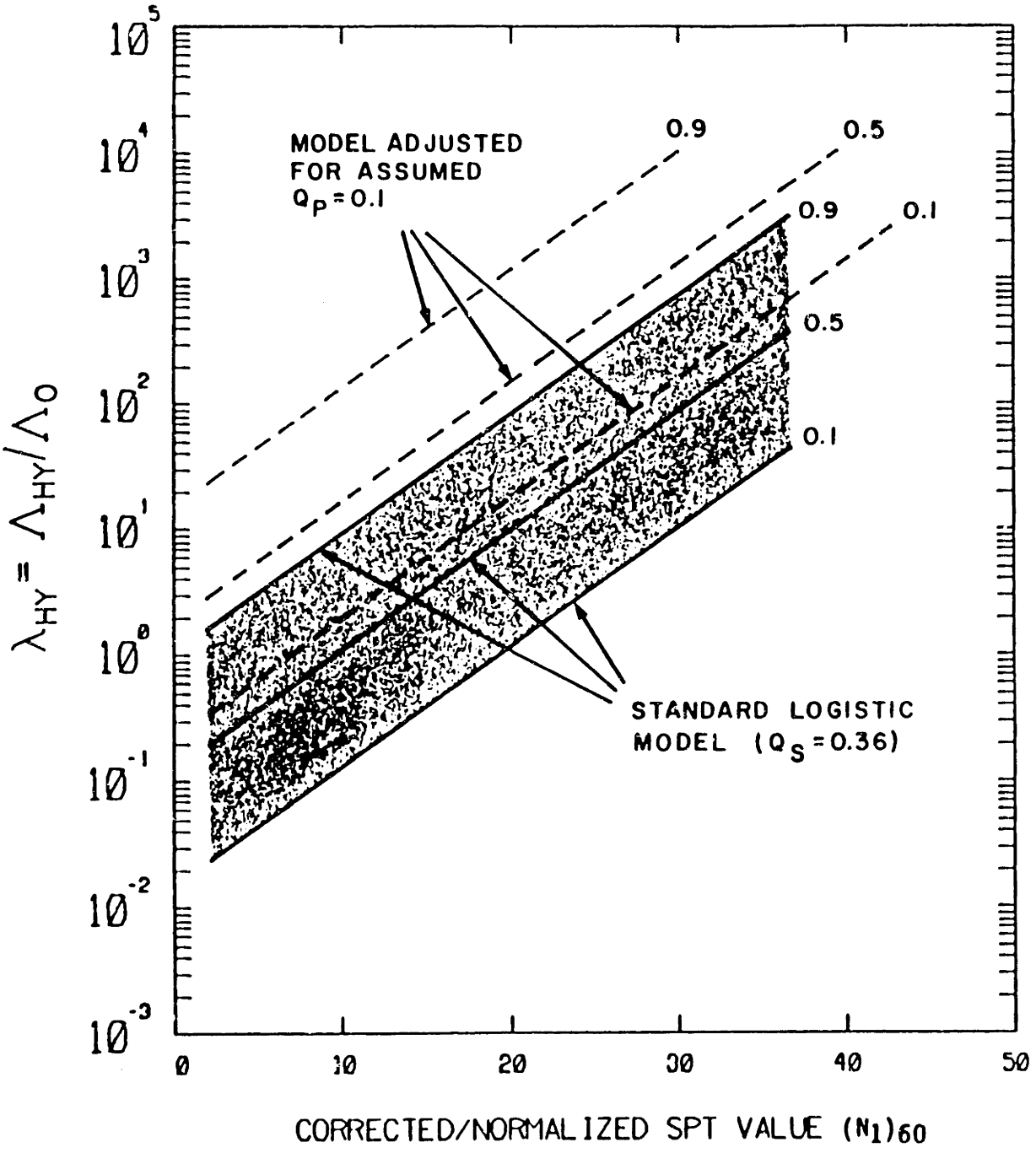
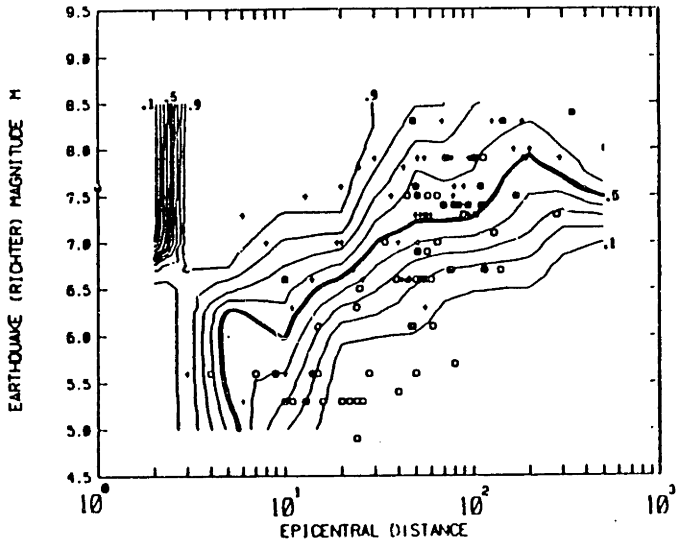
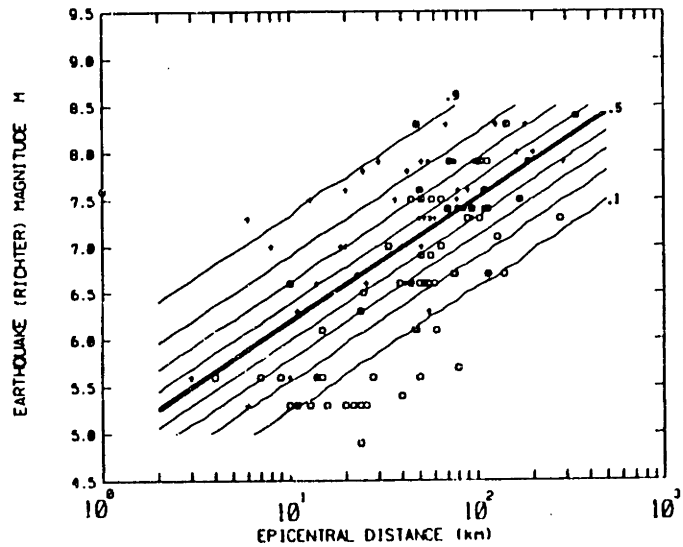


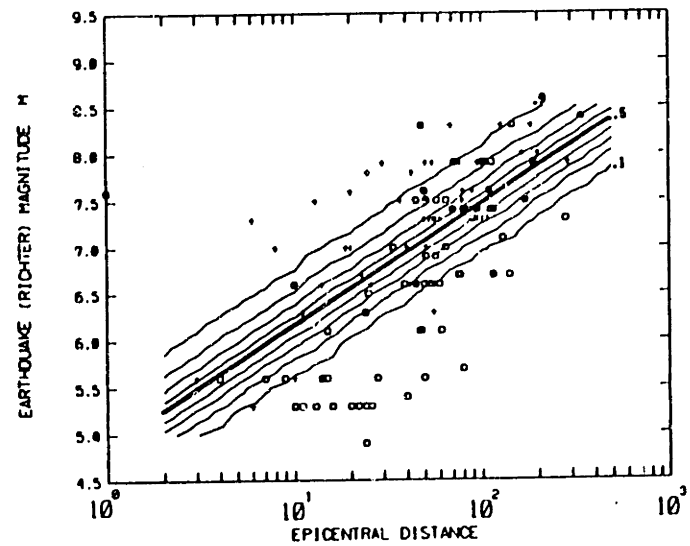
Fig. 6.12 Comparison of 0.1, 0.5, and 0.9 contour lines of equal probability of liquefaction for the standard logistic model and for the model corrected for the effects of choice based sampling (assumed Q_P = 0.1.). Modified Davis-Berrill (hypocentral) model.



(a) Kernel estimate of $Q_S(M,R)$. CVL optimum fixed size kernel: h=0.12, γ=3.1 (in M vs log R_{EP} space as shown)

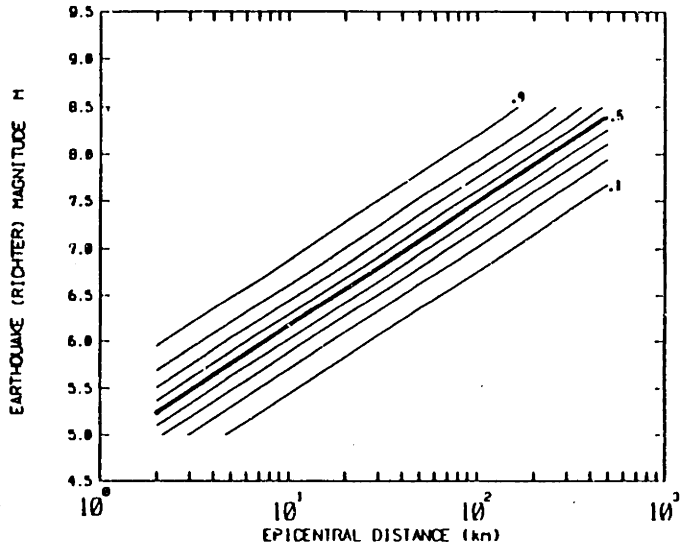


(b) Initial estimate of $Q_P(M,R)$ from standard logistic model.

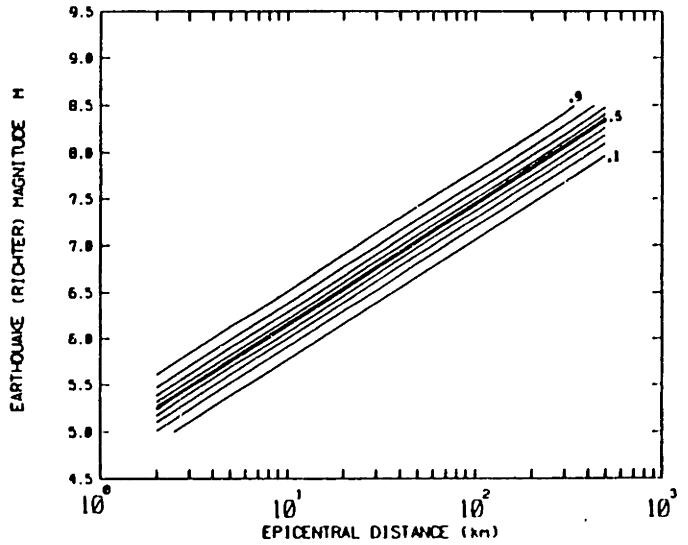


(c) $Q_P(M,R)$ at convergence of M&R bias correction procedure.

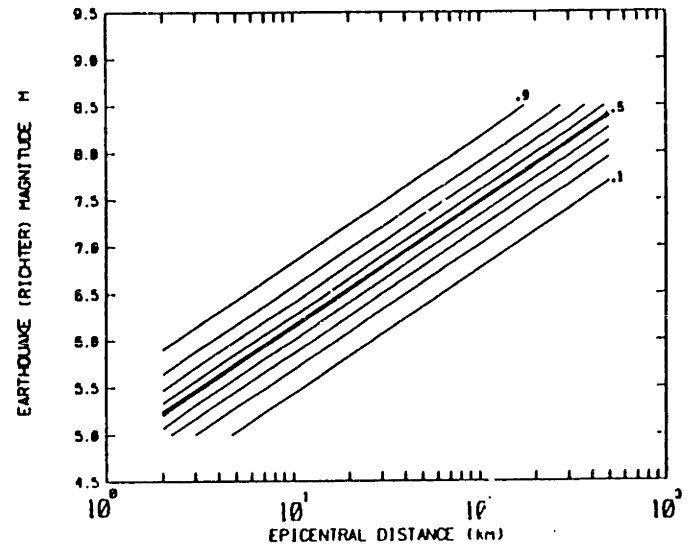
Fig. 6.13 Estimates of $Q_S(M,R)$ and $Q_P(M,R)$ for illustrating M&R bias correction algorithm.



(a) Standard logistic general M&R model



(b) General M&R bias-corrected logistic model



(c) Davis-Berrill epicentral logistic model

Fig. 6.14 Comparison of contour lines of equal probability of liquefaction in M vs log (R_{EP}) space for various logistic models. Contours for $(N_1)_{60}$ = 10 and $\bar{\sigma}_v$ = 0.7 kg/cm^2.

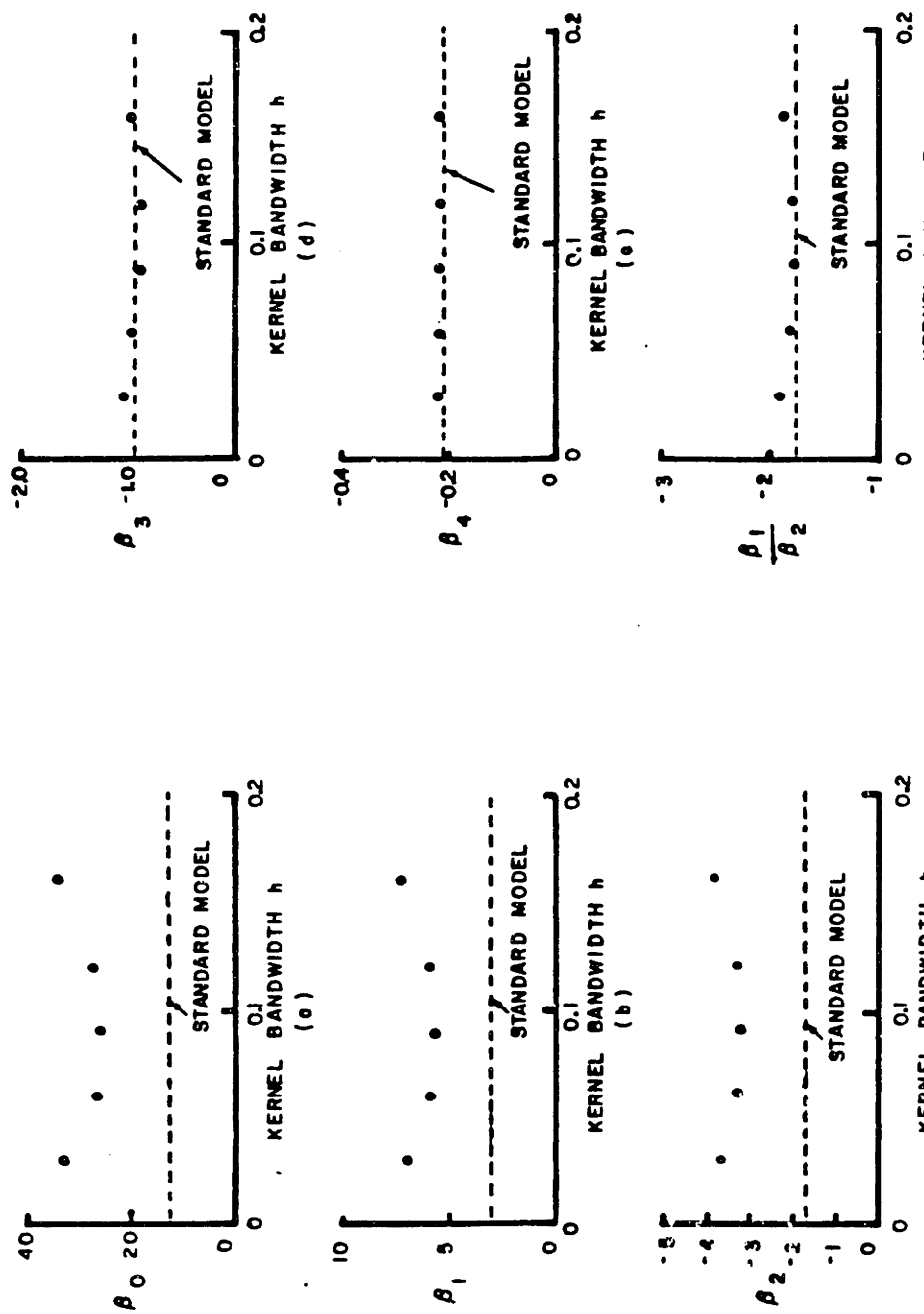
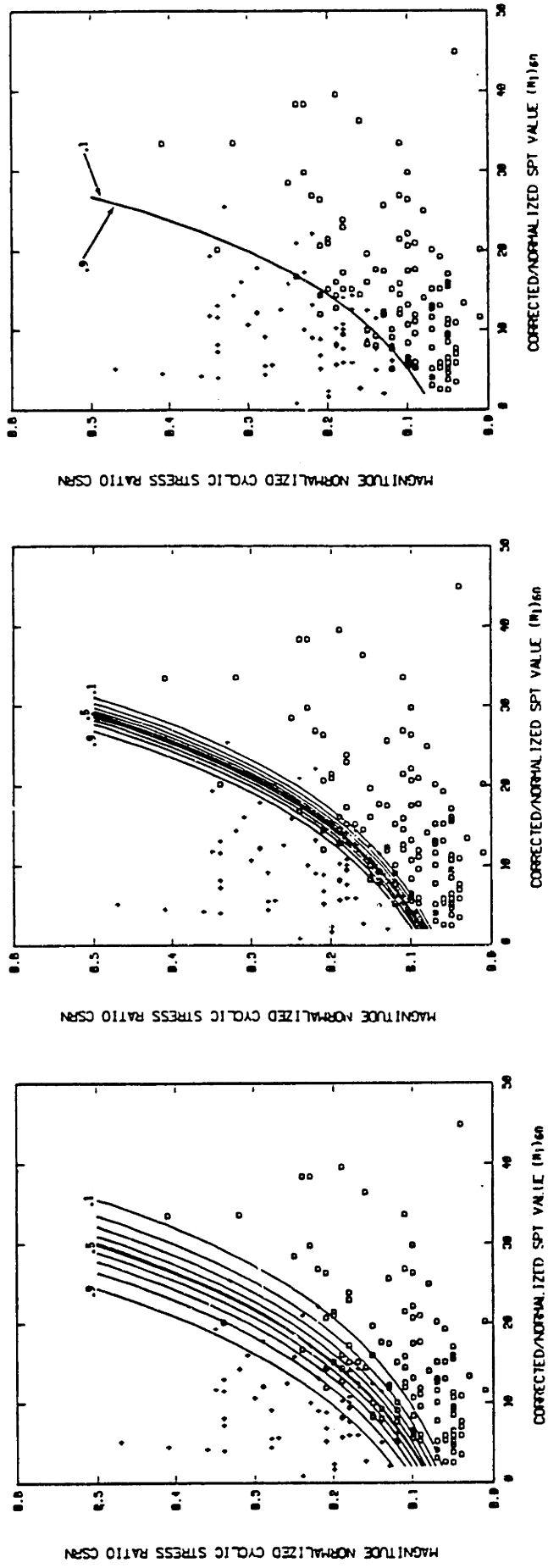


Fig. 6.15 Variation of logistic regression coefficients as a function of kernel bandwidth h (in units of log (R_{EP}(km))). Bandwidth ratio γ = 3.1.

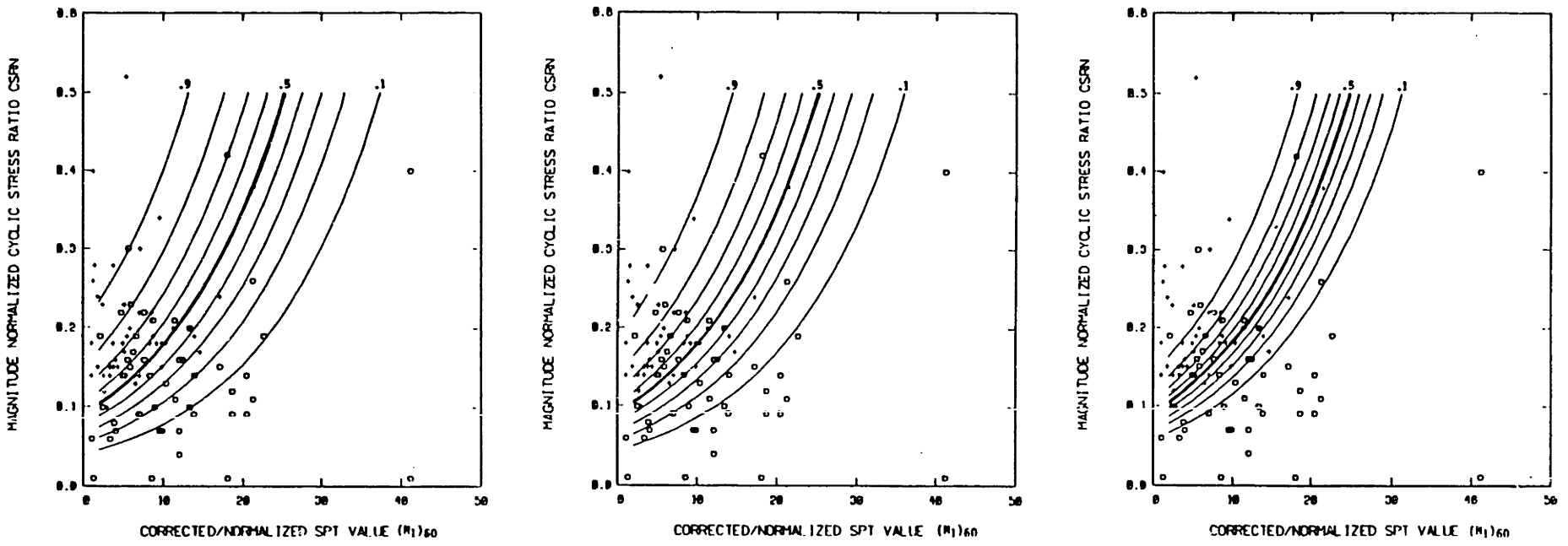


(a) $\sigma_1 = 0$
$V_2 = 0$

(b) $\sigma_1 = 0.25$
$V_2 = 0.1$

(c) $\sigma_1 = 0.5$
$V_2 = 0.2$

Fig. 6.16 Comparison of contours of equal probability of liquefaction P^* for various assumed levels of errors in variables. Seed-Idriss model - clean sand data. (EIV likelihood function evaluated using 7x7 hermite integration scheme.)



(a) $\sigma_1 = 0$
$V_2 = 0$

(b) $\sigma_1 = 0.25$
$V_2 = 0.1$

(c) $\sigma_1 = 0.5$
$V_2 = 0.2$

Fig. 6.17 Comparison of contours of equal probability of liquefaction P* for various assumed levels of errors in variables. Seed-Idriss model - silty sand data. (EIV likelihood function evaluated using 7x7 hermite integration scheme.)

CHAPTER 7

IMPLEMENTATION OF LOGISTIC MODELS IN RISK ANALYSIS

7.1 Introduction

Several logistic models for estimating the conditional probability of liquefaction were obtained in Chapter 4. In Chapter 5, a nonparametric formulation was used to verify the validity of the mathematical form of the logistic models. Modifications of standard logistic regression were implemented in Chapter 6 to investigate the effects of errors and biases in the data. Although the more elaborate models have features of interest, they are not recommended for use in seismic risk analysis, due largely to the limitations of the data base. Of all the models investigated, the following four were considered to be adequately supported by the data and hence to be useful for routine application:

- Seed-Idriss Clean/Silty Sand Model: Based on using the cyclic stress ratio CSRN, which is a quantity similar to that originally defined by Seed and Idriss (1971). The model differentiates between clean and silty sand.
- Seed-Idriss Base Model: Same as the previous model, except that there is no differentiation between clean and silty sand. The context in which this model is useful will be described later.
- Davis-Berrill Model: Based on an earthquake load parameter Λ_{EP}, which is similar in structure to a parameter proposed by Davis and Berrill (1981). Λ_{EP} is a function of earthquake magnitude and epicentral distance.

- Modified Davis-Berrill Model: Same as the Davis-Berrill model, except that hypocentral instead of epicentral distance is used in calculating the load parameter Λ_{HY}.

The corrected/normalized SPT $(N_1)_{60}$ value is used in each of the above models to characterize soil liquefaction resistance. It should be noted that the Davis-Berrill models are insensitive to the effects of fines content, as previously discussed in Chapter 4. The fitted logistic equations for these models are shown in Table 7.1.

Section 7.2 discusses the applicability of the above models for various problems, and Section 7.3 makes some remarks on site characterization. Comparisons of the logistic models with the deterministic procedures of Seed et al. (1984) and Davis and Berrill (1982, 1985) are presented in Section 7.4. Calculations for a typical site are performed in Section 7.5 to illustrate the features of each of the logistic models. An interpretation of the logistic models in the context of plots of the type proposed by Kuribayashi and Tatsuoka (1975) is discussed in Section 7.6.

7.2 Model Selection in Applications

Each of the logistic models given in Table 7.1 has an appropriate domain of application. This section presents the issues that should be considered in selecting a model for use in various situations. Only general guidelines are presented and it is recognized that the circumstances surrounding the actual choice are in many cases unique. A schematic of the decisions one must make in model selection is shown in Fig. 7.1.

The first issue in model selection is the decision between using a

local A&D or global M&R formulation. The Seed-Idriss models, which represent the A&D approach, are based on the physical concept that liquefaction is caused by cyclic shearing stresses due to earthquake shaking. The Davis-Berrill models, representing the M&R formulation, are derived from considerations of earthquake energy dissipation that give rise to excess pore pressures causing liquefaction. [Not all M&R formulations are derived in this way. See Chapter 4 for more details.] Both types of model fit the data well, though the Seed-Idriss approach has somewhat better goodness-of-fit statistics. However, the Davis-Berrill models are considered to be "closer" to the data or more "empirical" in a positive sense. This is because the magnitude and distance measures in the data catalog are believed to be less subject to potential biases (due to interpretation or judgement) than is the peak site acceleration used for the Seed-Idriss models; see Section 4.7.

On a methodological level, the Davis-Berrill models are conceptually more appropriate for use in seismic risk analysis. For these models, the probability of liquefaction is directly related to M and R, whereas in the Seed-Idriss models, the liquefaction probability is a function of cyclic stress ratio which must be calculated based on peak acceleration estimated through an attenuation relationship. In addition, the uncertainty in the estimated value of acceleration needs to be incorporated in the A&D approach. There are theoretical reasons within the context of regression analysis why such a procedure is strictly incorrect (Heidari, 1986), though the effects on the calculated probabilities of liquefaction may be small.

The Davis-Berrill load parameter implicitly incorporates an attenuation relationship. Though the Davis-Berrill formulation fits the

data fairly well, it is recognized that the model is fitted to data consisting primarily of case studies from California and Japan. Thus, the Davis-Berrill model is "locked in" to those geographic regions while the Seed-Idriss model may be more transportable to other regions with different attenuation relationships. However, caution should be exercised in making such a conclusion because the Seed-Idriss procedure is also limited by the data. The cyclic stress ratio CSRN is directly proportional to the peak ground acceleration for the site, but does not account for the frequency content of the strong motion record, and hence is an incomplete characterization of the earthquake load. It is well known that earthquakes of different magnitudes and at different distances have different frequency content, even if they produce the same peak acceleration.

In view of the above considerations, it is recommended that either the Seed-Idriss or Davis-Berrill models can be used in regions where the attenuation of earthquake motions are similar to those encountered in California and/or Japan. Other differences aside, the Davis-Berrill models are favored in terms of methodological simplicity and correctness. However, in regions where attenuation is considered to be vastly different from that in California or Japan, the adaptability of the Seed-Idriss models makes these better suited to such applications.

Assuming that it has been decided to employ the Seed-Idriss approach, the next issue involves a choice between using a model that distinguishes between clean and silty sand or a model that does not. The resolution of this issue largely involves the question of whether the time and expense of acquiring fines content data is worth the refinement

of results obtained. For example, if the analysis involves a major facility, fines content data can usually be acquired as a matter of routine testing of samples obtained during the soil investigation. In such a case, use of the model which distinguishes between clean and silty sand is certainly recommended. On the other hand, if one is interested in large-scale mapping of liquefaction hazards and is faced with assessing soil conditions from maps of surficial geology, then acquiring data to classify sands as clean or silty might represent a substantial cost. An intermediate approach to this problem would be to assess the relative frequency of occurrence of clean and silty sands within the region to be mapped. Then the relative frequencies can be incorporated into the Seed-Idriss Clean/Silty Sand model. However, as will be shown by the example calculations in Section 7.5, uncertainty in seismicity and attenuation tends to mask the effects of soil gradation, so that even this intermediate approach may not be worth the effort. For most large-scale mapping projects, use of the Seed/Idriss Base model will be adequate.

If the Davis-Berrill approach is to be used in analysis, then a decision has to be made regarding whether to employ the epicentral or hypocentral formulation. This depends largely on the state of knowledge regarding earthquake focal depths and whether the analysis is being performed for the near field or far field of the earthquake. In the case where the earthquake sources are located at epicentral distances greater than two times the focal depth, the epicentral formulation can be used without relying on information regarding focal depths. However, for closer distances to the source, the hypocentral model is preferable and is necessary to assure reasonable estimates of the probability of

liquefaction.

7.3 Notes on Site Characterization

Site characterization refers to the problem of how to choose a representative SPT N-value for a soil deposit. In Chapter 3, this problem was discussed in the context of assembling the data catalog used in this study. In applying models fitted through statistical analysis, it is logical to use the same site characterization procedure as employed in compiling the data. This procedure consists of:

- Defining each boring drilled as the basic unit of analysis.
- Choosing the minimum SPT N_1-value measured in the liquefiable stratum in each boring. Judgement should be employed to exclude questionably low N_1-values that may be due to mistakes in the drilling or testing procedures.

Note that the values of N_1 should be converted to $(N_1)_{60}$. Alternatively, a simple adjustment can be made to the logistic models (Table 7.1) to reflect the hammer energy standard used on the project, e.g. if the energy ratio used is ER = 55% then the logit coefficient of $(N_1)_{60}$ should be multiplied by 60/55 and $(N_1)_{55}$ can be used instead of $(N_1)_{60}$. In analysis, it is recommended to calculate the probability of liquefaction at each boring location, and then appropriately interpolate to obtain values between the borings.

7.4 Comparison with Deterministic Criteria

Figures 7.2 and 7.3 show contours of the conditional liquefaction probability from the Seed-Idriss clean/silty sand model, superimposed on the liquefaction criteria recently proposed by Seed et al. (1984). If one considers the contour for P = 0.5 to be a reasonable discriminant

line, then the logistic model and the Seed et al. criteria are in overall good agreement. However, there are several differences which should be noted:

1) In the logistic model, clean sand is defined as sand with fines content FC < 12%, whereas the Seed et al. criteria defines it as FC < 5%.

2) The Seed et al. criteria imply that as FC increases, there is a continuous increase in liquefaction resistance. Statistical analysis indicates that the current data does not support this hypothesis.

3) For clean sand (Fig. 7.2), the Seed et al. discriminant line goes through the origin, while the logistic model contours have positive intercept on the CSRN axis. Thus, the Seed et al. criteria is very conservative for low values of $(N_1)_{60}$.

Several results have been presented in Chapters 4 and 5 that indicate the features of the logistic model to be more appropriate than the Seed et al. criteria.

In Fig. 7.4, the Davis-Berrill logistic model is compared with the discriminant analysis results obtained by Davis and Berrill (1982, 1985). The discriminant lines display a concave downward curvature in this figure, but would be linear in a plot where $(N_1)_{60}$ is scaled logarithmically. [This reflects the way in which linear discriminant analysis is used by Davis and Berrill (1982,1985).] However, their 1982 discriminant line agrees fairly well with the P = 0.5 logistic regression contour for values of $(N_1)_{60} < 20$, due to the fact that the bulk of the data have values of $(N_1)_{60}$ in this range. However, their 1985 line agrees less well with the logistic model. The difference between their 1982 and

1985 lines results primarily from an increased amount of (and therefore different) data in their 1985 compilation. [See Section 2.4.4 for a discussion of the sensitivity of discriminant analysis to changes in the data base.] An additional refinement to include effects of material damping of earthquake waves is contained in the 1985 Davis and Berrill procedure. However, as noted by Davis and Berrill (1985), adjustments from this refinement are relatively small.

In summary, it is expected that the results of risk analysis using the proposed logistic models will be generally in agreement with conclusions that may be drawn from deterministic analyses using the Seed et al. (1984) criteria. Less agreement is expected with analyses using the most recent procedures recommended Davis and Berrill (1985), especially for sites with $(N_1)_{60} > 20$.

7.5 Example Calculations for a Typical Site

In this section, calculations performed for a typical site are presented to illustrate the different features of the Seed-Idriss and Davis-Berrill logistic models. The conditional probability of liquefaction is calculated for various combinations of earthquake magnitude (M) and epicentral distance (R_{EP}), and the results are plotted as contours of equal liquefaction probability in the M versus log(R_{EP}) plane. This provides a common format to compare the implications of each of the models.

The calculations were performed for earthquake magnitudes ranging from 5.0 to 8.5 and for epicentral distances from 2 km to 500 km. It should be noted that the majority of the data used to fit the logistic models occur for magnitudes between 6.0 and 8.0 and for epicentral

distances between 10 km to 150 km, as shown in the histograms (Figures 3.5 through 3.6) presented in Chapter 3. Thus, to a certain extent, calculations outside the ranges where the data is concentrated represent extrapolations that should be considered less reliable when interpreting the results of this section.

7.5.1 Assumptions

The soil profile assumed for the calculations consists of a deep sand stratum with the water table at a depth of 2m. The saturated unit weight of the soil is assumed to be 1.8 kg/cm^3 (110 PCF) and the critical liquefaction depth is taken to be at 6m. This results in total and effective stresses of σ_v = 1.0 kg/cm^2 and $\bar{\sigma}_v$ = 0.7 kg/cm^2 at that depth. A value of $(N_1)_{60}$ = 10 is used in the calculations.

In order to implement the Seed-Idriss models, assumptions for the attenuation relationship are also necessary. Calculations were performed using the relationship from Iwasaki et al. (1978):

$$\frac{a}{g} = \frac{0.0188\ 10^{0.302 \cdot M}}{(R_{EP})^{0.8}} \qquad (7.1)$$

and that from Kawashima et al. (1984) for soft alluvium or reclaimed land:

$$\frac{a}{g} = \frac{0.4109\ 10^{0.262 \cdot M}}{(R_{EP} + 30)^{1.208}} \qquad (7.2)$$

The principal difference between the above two attenuation equations is in the denominator term. Note that when $R_{EP} \rightarrow 0$, a/g approaches infinity in Eqn. 7.1, whereas a/g approaches a finite value in Eqn. 7.2. Thus the calculated acceleration levels from Eqns. 7.1 and 7.2 will differ most significantly in the near field of the earthquake source. These two

attentuation relationships were selected for the example calculations because they are formulated in terms of epicentral distance. In addition, they are based upon data taken in the same geographic region (i.e., Japan) from which most of the observations of liquefaction or non-liquefaction originate. Other relationships such as that of Joyner and Boore (1981), which is formulated in terms of closest distance to the surface projection of the fault rupture, are less convenient for the purposes of the comparisons presented here. Uncertainty in the accelerations obtained using the attenuation relationships was incorporated by assuming a log-normal dispersion of the error ε about the estimate of a/g, with $\sigma_\varepsilon[\ln(a/g)] = 0.5$.

7.5.2 Seed-Idriss Models

Fig. 7.5 shows the results of calculations using the Seed-Idriss base model (which does not distinguish between clean and silty sand) combined with the Iwasaki et al. (1978) attentuation equation. The resulting contours of equal liquefaction probability are straight lines in the M vs. $\log(R_{EP})$ plane. [The small oscillations of the lines are due to the method of contouring.] In Fig. 7.5(a), the contours are obtained assuming no uncertainty in the attenuation relationship. When uncertainty in attenuation is introduced as shown in Fig. 7.5(b), the dispersion of the equi-probability lines about the P = 0.5 contour increases.

A similar set of contours is shown for the Seed-Idriss Clean/Silty Sand model in Fig. 7.6, where no uncertainty in attenuation is assumed, and in Fig. 7.7, where this uncertainty is incorporated. The clean and silty sand contours for P = 0.1, 0.5 and 0.9 are superimposed in Fig. 7.8

for comparison. As is consistent with the features of this model described in Chapter 4, the P = 0.5 contour for silty sand is slightly higher than that for the clean sand, but the dispersion of the silty sand contours is greater. This indicates that silty sands are on the average less susceptible to liquefaction, but the prediction of liquefaction is more uncertain for silty than for clean sands. However, as shown in Fig. 7.8(b), the introduction of attenuation uncertainty in conditional probability calculations tends to obscure the differences in the contours of the clean and silty sands. This is consistent with the observation (in Chapter 4) that the Davis-Berrill models are insensitive to soil gradation charateristics. Since the Davis-Berrill models intrinsically incorporate uncertainty of attenuation, the refinements of soil gradation are consequently obscured.

A practical conclusion that can be drawn from the above result is that it may not be worthwhile to try to distinguish between clean and silty sands in risk analysis where there is a great deal of uncertainty in attenuation or seismicity in general. For example, in regional mapping of liquefaction risk, considerable effort and expense may be involved to acquire data on whether particular deposits of sand are silty or clean. It is recommended that the level of knowledge regarding regional seismicity be considered in decisions as to whether to acquire this data or not. If the knowledge of seismicity is relatively meager, then the effort and expense may not be justified.

The straight line contours in Figures 7.5 through 7.8 are a result of the Iwasaki et al. (1978) attenuation relationship used in the calculations. An alternate set of results obtained using the Kawashima et al. (1984) equation is shown in Fig. 7.9. Thus, the results of risk

analysis using the Seed-Idriss models is a function of the assumed attenuation of peak acceleration with distance. This is an advantage in that the Seed-Idriss model can be adapted to regions with attenuations different from those that constitute the data base used for model fitting.

7.5.3 Davis-Berrill Models

The Davis-Berrill models are relatively easier to implement because the probability of liquefaction is directly related to earthquake magnitude and site-to-source distance, and thus the assumption of an attenuation relationship is unnecessary. This is both an advantage and a disadvantage as discussed previously in Section 7.2. Consideration of the uncertainty of attenuation is also intrinsically incorporated into the models. Fig. 7.10(a) shows the contours of equal liquefaction probability obtained from the epicentral Davis-Berrill model. The contours are straight lines whose slope is fixed by the definition of the load parameter Λ_{EP}. For purposes of comparison, the results from Seed-Idriss model combined with the Iwasaki et al. (1978) attenuation relationship is shown in Fig. 7.10(b). The slopes of the contour lines for the two models are significantly different, with the Seed-Idriss model predicting lower probabilities of liquefaction for large epicentral distances. As noted in Section 7.2, the Davis-Berrill model implicitly contains an attenuation law. Thus, one should expect calculations using the Seed-Idriss and Davis-Berrill models to be systematically different, and that this difference would be dependent on the attenuation equation used in conjunction with the Seed-Idriss model.

Equi-probability contours of a concave upward shape can be obtained

using the Modified Davis-Berrill model. Since the load parameter Λ_{HY} is calculated using hypocentral distance, which can be expressed as

$$R_{HY} = (R_{EP}^2 + H^2)^{1/2} \tag{7.3}$$

where H is the focal depth, the contours are non-linear in the M versus log(R_{EP}) plane. Contours for various values of H are shown in Fig. 7.11.

7.6 Commentary on the Kuribayashi-Tatsuoka Plot

In 1975, Kuribayashi and Tatsuoka published a paper in which they compiled historical data on earthquake-induced liquefaction in Japan dating back to the 1800's. For each earthquake, they plotted (on a graph of magnitude versus epicentral distance) the distance to the furthest site from the epicenter at which liquefaction was observed. An extremal bound to the data was proposed as a simple liquefaction criterion. An updated version of this graph is shown in Fig. 7.12 with additional data plotted by Youd (1977), Davis and Berrill (1983) and Seed et al. (1984). An alternative extremal bound to the data proposed by Seed et al. (1984) and a criterion developed by Youd and Perkins (1978) are also shown.

There are several problems associated with the Kuribayashi-Tatsuoka plot. The most obvious is that liquefaction resistance of the soil is not taken into account for any of the sites. Another deficiency is that the data points tend to be biased towards conservatism. Many of the data points, particularly those at large distances, may represent extreme occurrences or anomalies. For example, an occurrence of liquefaction at a large epicentral distance may be due to the local amplification of ground motion rather than reflecting a global phenomenon as implied by the Kuribayashi-Tatsuoka plot. Despite these problems, a commentary on

the plot is attempted in light of the calculations of Section 7.5. This commentary is not motivated by a desire to justify the logistic models by comparing them to the data plotted by Kuribayashi and Tatsuoka in a format that is clearly inappropriate for the purpose of liquefaction risk evaluation. Rather, the present comments aim at indicating that there are better alternatives to plots of the Kuribayashi-Tatsuoka type.

The plot of Fig. 7.12 has formed the basis of the Youd and Perkins (1978) procedure which has been adopted by the USGS for regional mapping of liquefaction risk. The Kuribayashi-Tatsuoka plot has also been used by Davis and Berrill (1982, 1983) and Seed et al. (1984) to argue for support of their methods of liquefaction analysis. It is of interest to note that the slope of the line proposed by Kuribayashi and Tatsuoka (Fig. 7.12) is practically the same as the slope of the equi-probabilility contour lines of the Davis-Berrill logistic model shown in Fig. 7.10(a). The slope of the line proposed by Seed et al. (Fig. 7.12) approximately equals the slope of the contours in Fig. 7.10(b) and Figs. 7.5 through 7.8, which show calculations using the Seed-Idriss logistic models combined with the Iwasaki et al. (1978) attenuation relationship.

The interpretation of the data by the writer is that they are more appropriately represented as shown by the shaded region in Fig. 7.12. In contrast to previous interpretations, the pattern is viewed to have a curvature, while others have tried to model the data in terms of straight lines. Based on the shape of the shaded region, it appears that the probability of liquefaction is very small for earthquake magnitudes less than 5.0. In Fig. 7.13, the pattern of the data is superimposed on the equi-probability contours obtained from two sets of calculations performed in Section 7.5 which were previously shown in Fig. 7.9(a) and

Fig. 7.11(b). The numerical values of probability of liquefaction are not important in this discussion. Rather, the emphasis is on the similar shapes of the contours and the pattern suggested by the points.

The conclusion offered is that, whether one believes that an extremal bound to the data should be represented as a straight line with a specific slope or as a curved contour, there exists a model that can be used to match that belief. The comparison in Fig. 7.13 suggests that the hypocentral Davis-Berrill model is more appropriate than the epicentral model. In employing the Seed-Idriss logistic models, attenuation relationships similar in form to those of Kawashima et al. (1984) would give results that comply with the data in Fig. 7.12 and Fig. 7.13. If the use of the Youd and Perkins (1978) procedure in liquefaction hazard mapping is to be continued, their liquefaction criterion (Fig. 7.12) should perhaps be updated to reflect the pattern of the more recent data. Rational ways to modify this criterion can be obtained through use of the logistic models.

7.7 Chapter Summary

In this Chapter, various aspects of the implementation of logistic models in liquefaction risk analysis have been considered. Criteria for selecting the appropriate model in various applications have been suggested, and the proper site characterization procedure has been described. Calculations were performed to illustrate important features of the logistic models and the differences among them.

From a comparison with data plotted in the format suggested by Kuribayashi and Tatsuoka (1975), it is concluded that the proposed logistic models can be used to obtain results that are consistent with

data obtained in a context different from this study. The results suggest that, if use of the Youd and Perkins (1978) method of liquefaction hazard analysis is to be continued, the procedure should at least be modified to use a more rational criterion that can be developed based on a logistic model.

The results of the Seed-Idriss logistic model in risk analysis will yield results consistent with the deterministic criteria proposed by Seed et al. (1984). However, the results obtained from employing the Davis-Berrill logistic models are expected to be different from those calculated by the most recent Davis and Berrill (1985) procedure.

Table 7.1

Logit Models Recommended for Use
in Liquefaction Risk Analysis

MODEL	FITTED LOGISTIC EQUATION
Seed-Idriss Model:	
Clean Sand	$\underline{x}^T\underline{\beta} = 16.447 + 6.4603 \ln(CSRN) - 0.39760 (N_1)_{60}$
Silty Sand	$\underline{x}^T\underline{\beta} = 6.4831 + 2.6854 \ln(CSRN) - 0.18190 (N_1)_{60}$
Seed-Idriss Base Model (No Differentiation Between Clean and Silty Sand)	$\underline{x}^T\underline{\beta} = 10.167 + 4.1933 \ln(CSRN) - 0.24375 (N_1)_{60}$
Davis-Berrill (Epicentral) Model	$\underline{x}^T\underline{\beta} = -12.922 + 0.87213 \ln(\Lambda_{EP}) - 0.21056 (N_1)_{60}$
Modified Davis-Berrill (Hypocentral) Model	$\underline{x}^T\underline{\beta} = -15.143 + 1.0837 \ln(\Lambda_{HY}) - 0.22656 (N_1)_{60}$

Conditional Probability of Liquefaction Expressed as: $P = 1/[1 + \exp(-\underline{x}^T\underline{\beta})]$

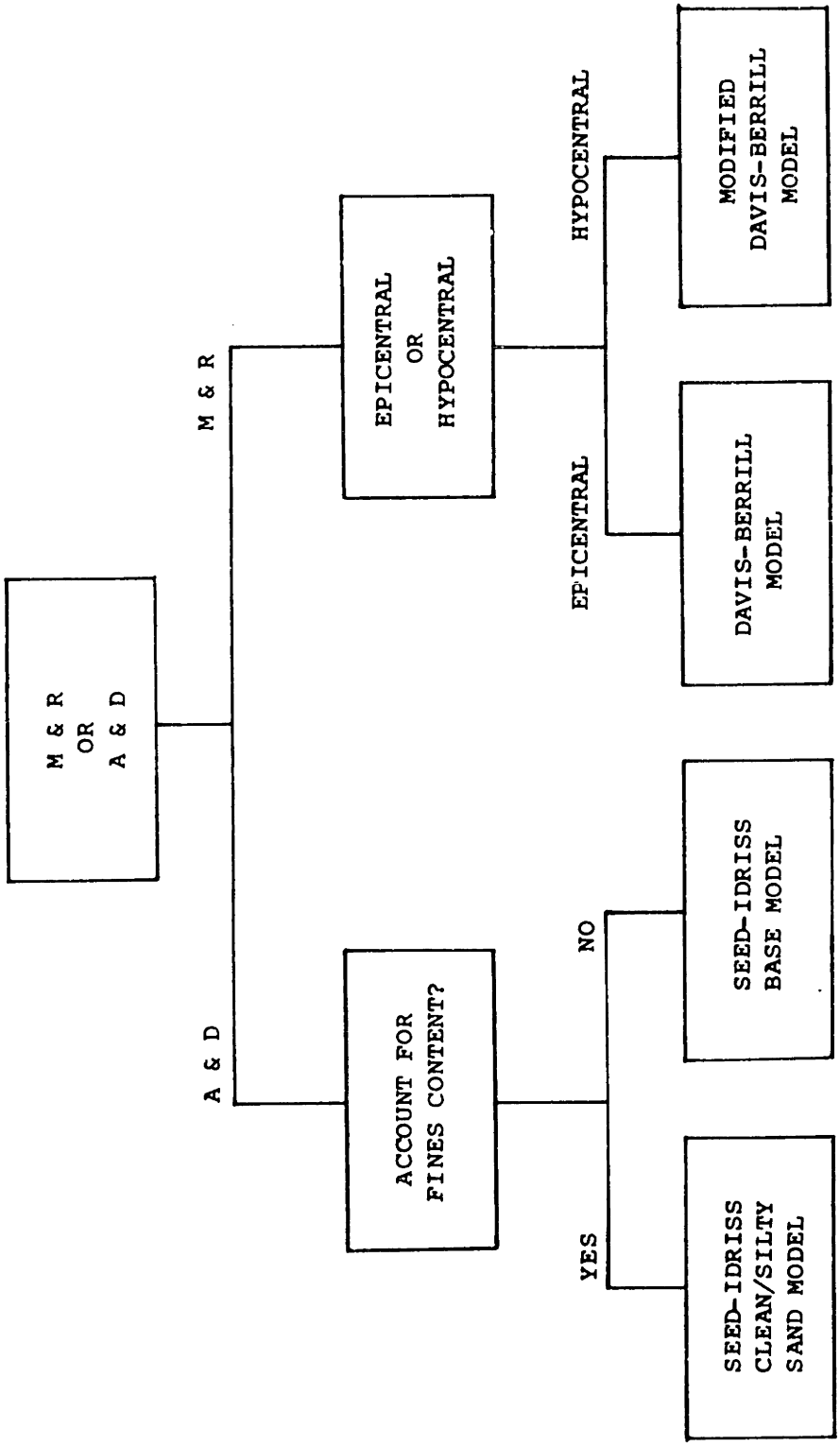


Fig. 7.1 Schematic of Decisions and Issues in Model Selection for Applications

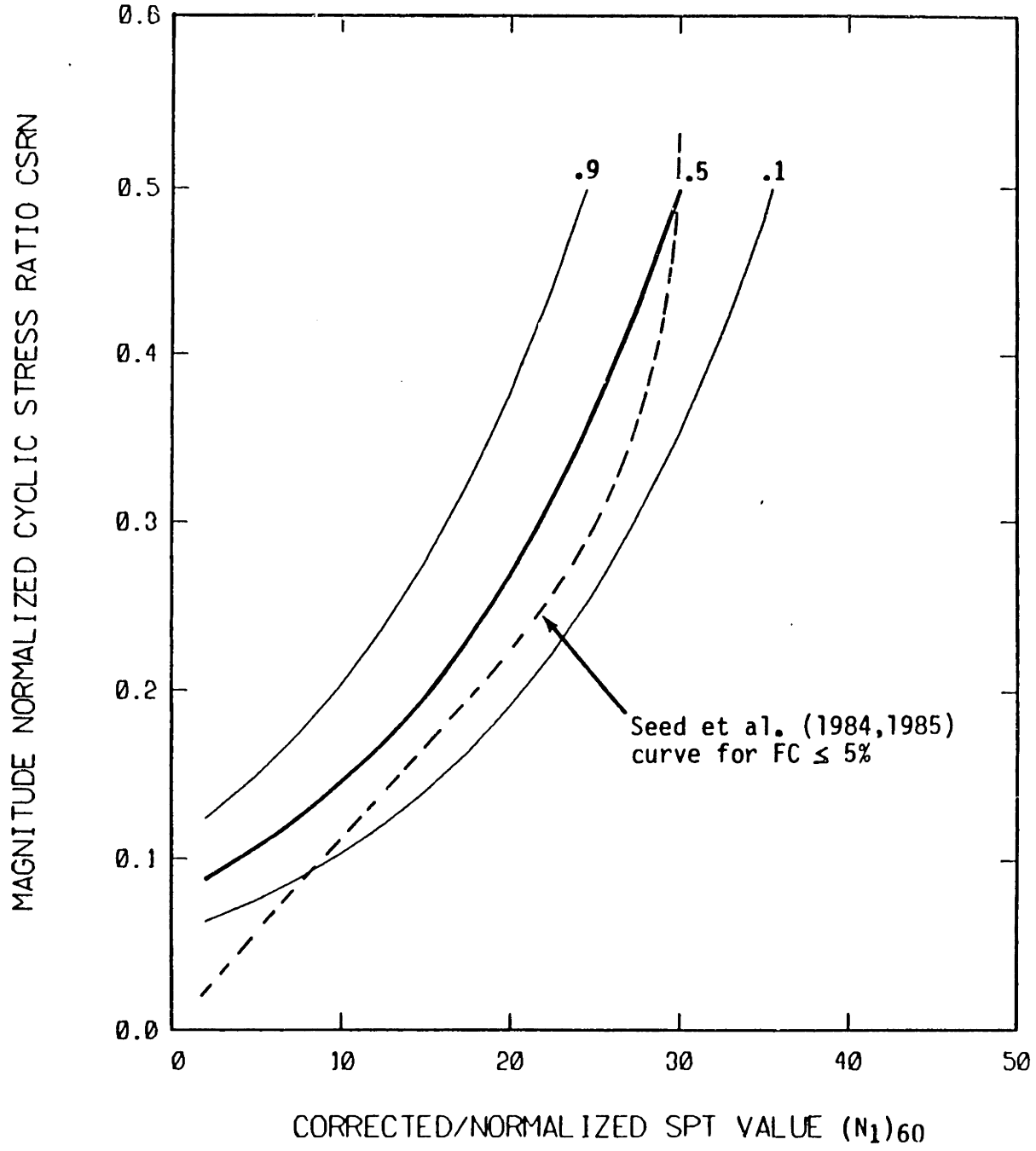


Fig. 7.2 Comparison of P = 0.1, 0.5, and 0.9 contours from logistic model for clean sands (FC < 12%) with deterministic line from Seed et al. (1984) for sands with fines content FC < 5%.

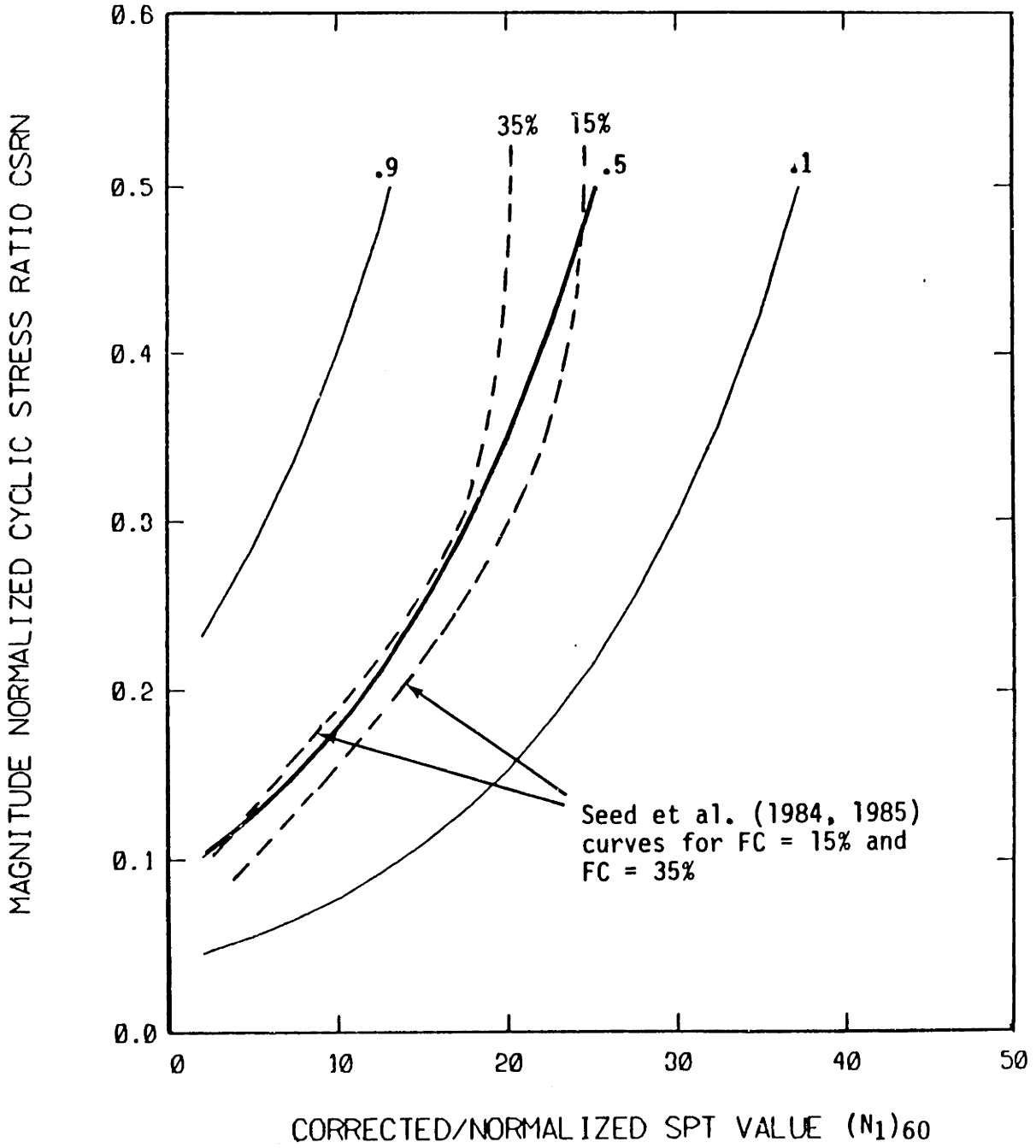


Fig. 7.3 Comparison of P = 0.1, 0.5, and 0.9 contours from logistic model for silty sands (FC > 12%) with deterministic line from Seed et al. (1984) for sands with fines content FC = 15% and FC = 35%.

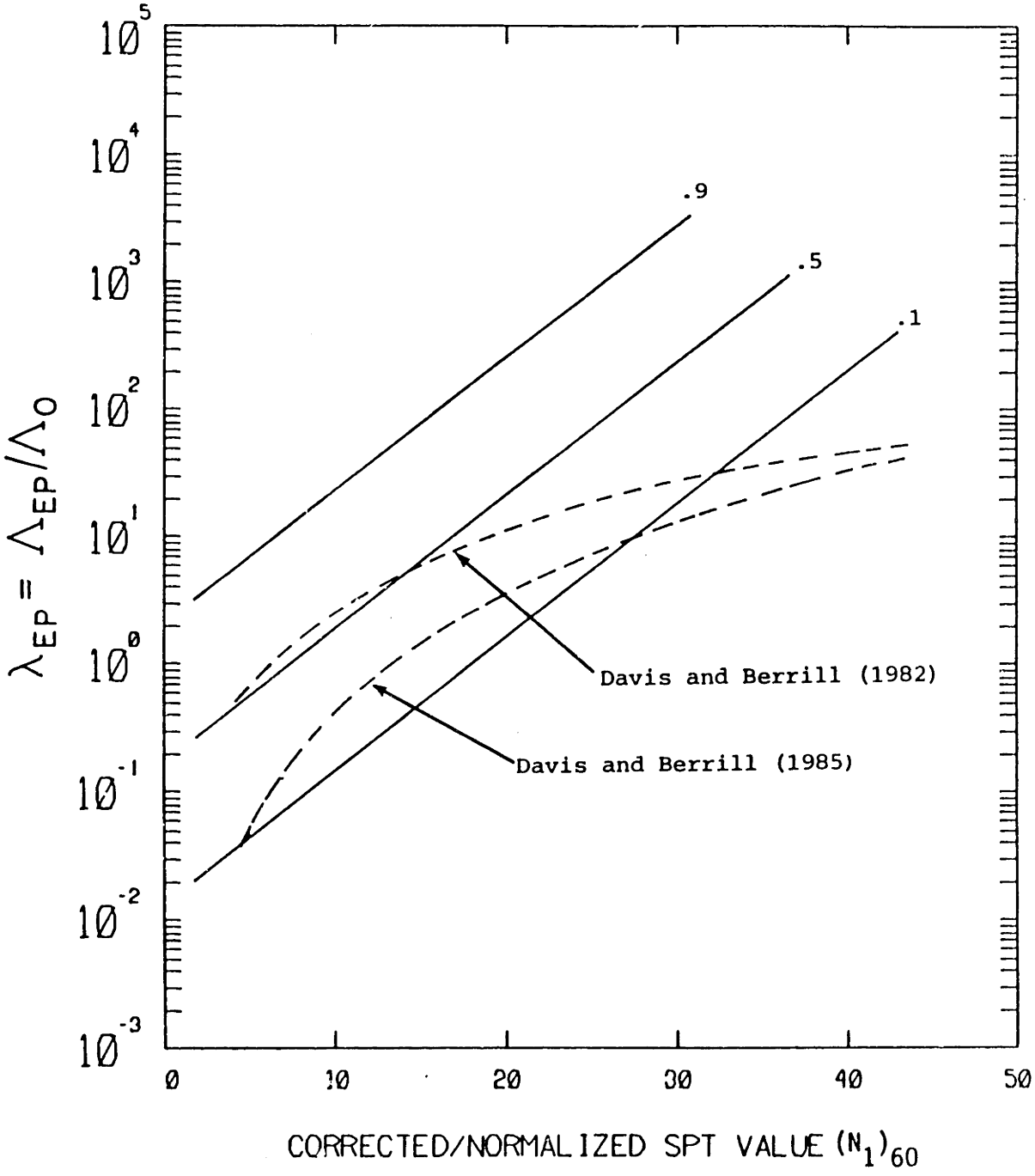
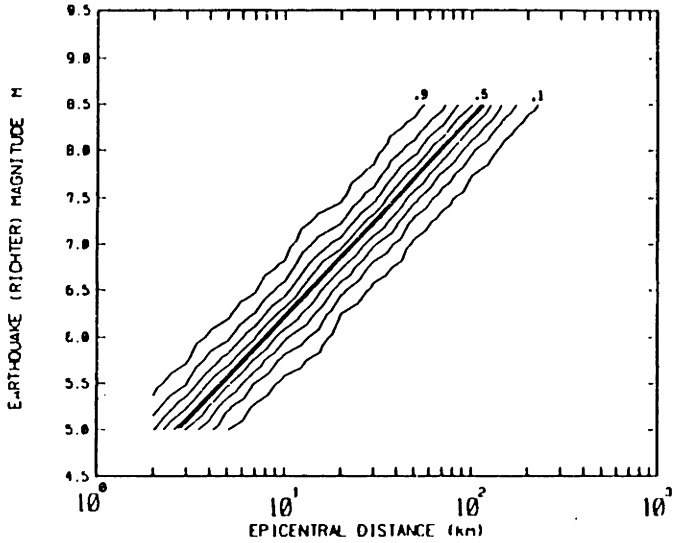
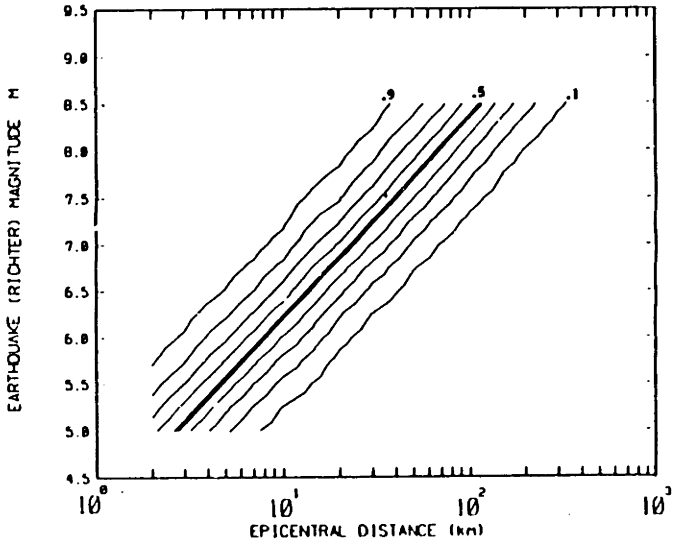


Fig. 7.4 Comparison of P = 0.1, 0.5, and 0.9 contours from logistic model with lines obtained by Davis and Berrill (1982, 1985) using discriminant analysis.

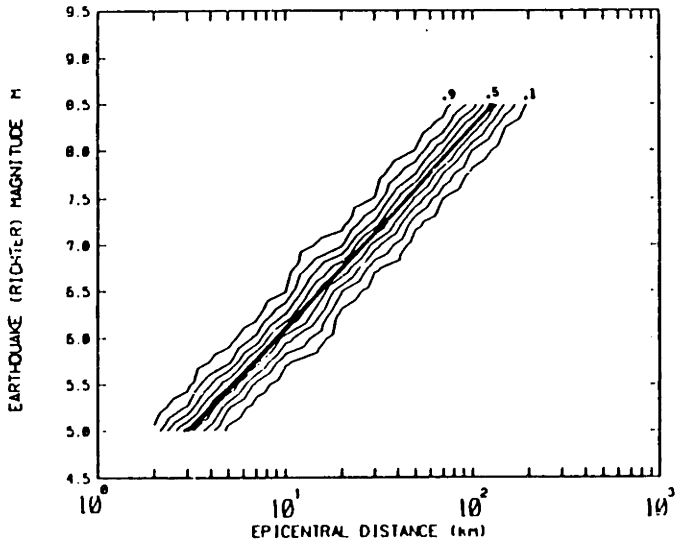


(a) No uncertainty in attenuation relationship

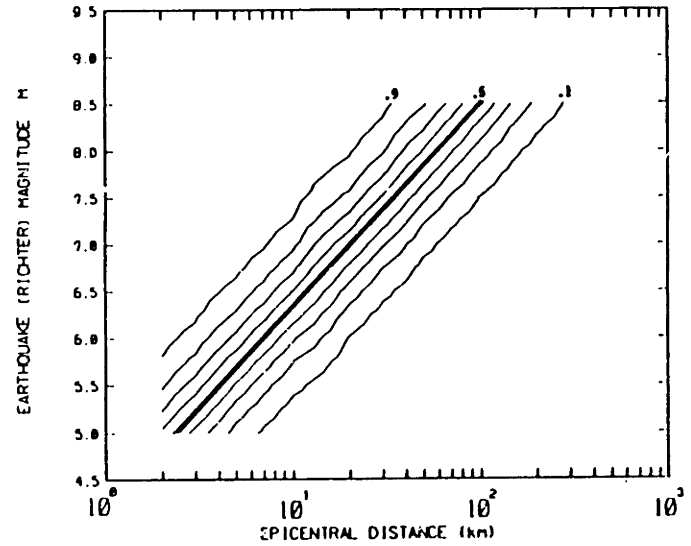


(b) Uncertainty included in attenuation relationship ($\sigma_\varepsilon[\ln(a/g)] = 0.5$)

Fig. 7.5 Comparison of lines of equal probability of liquefaction using the Seed-Idriss base model (no distinction between clean and silty sand). Calculations performed for a typical site using the Iwasaki et al. (1978) attenuation equation.

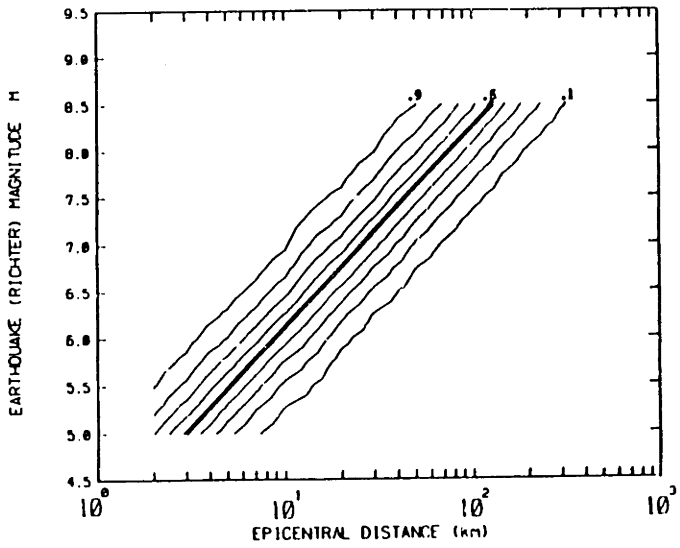


(a) Clean sand

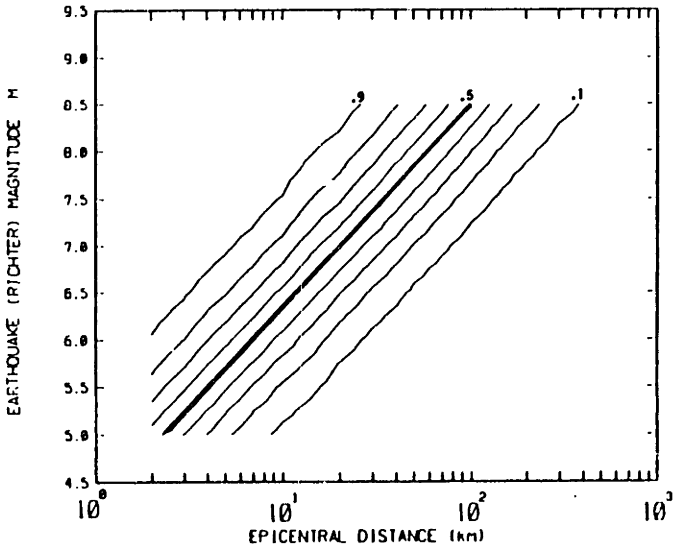


(b) Silty sand

Fig. 7.6 Lines of equal probability of liquefaction for a typical site, obtained using the Seed-Idriss model distinguishing between clean and silty sand. Attenuation relationship from Iwasaki et al. (1978). No uncertainty in attenuation assumed.

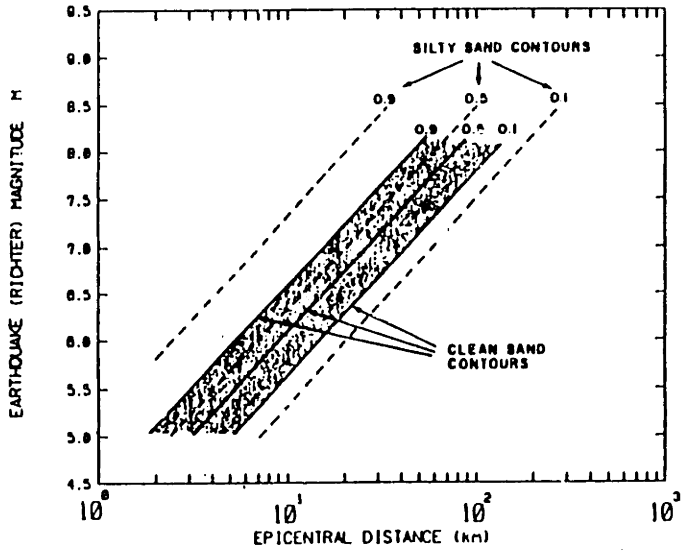


(a) Clean sand

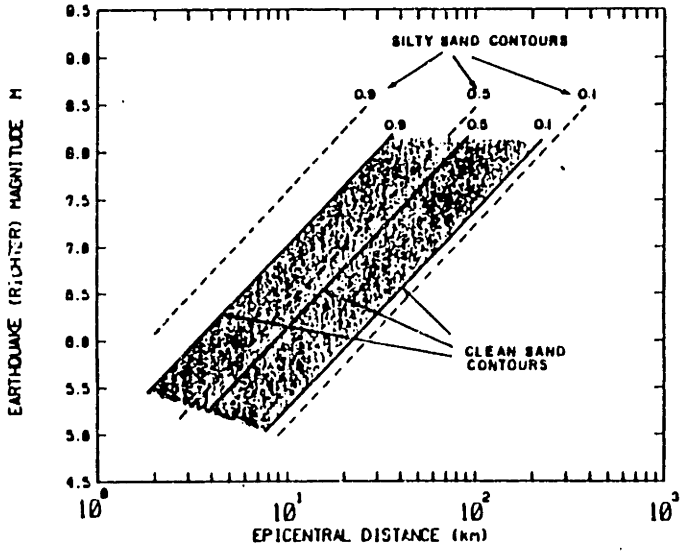


(b) Silty sand

Fig. 7.7 Lines of equal probability of liquefaction for a typical site, obtained using the Seed-Idriss model distinguishing between clean and silty sand. Attenuation relationship from Iwasaki et al. (1978). Uncertainty included in attentuation. ($\sigma_\varepsilon[\ln(a/g)] = 0.5$)

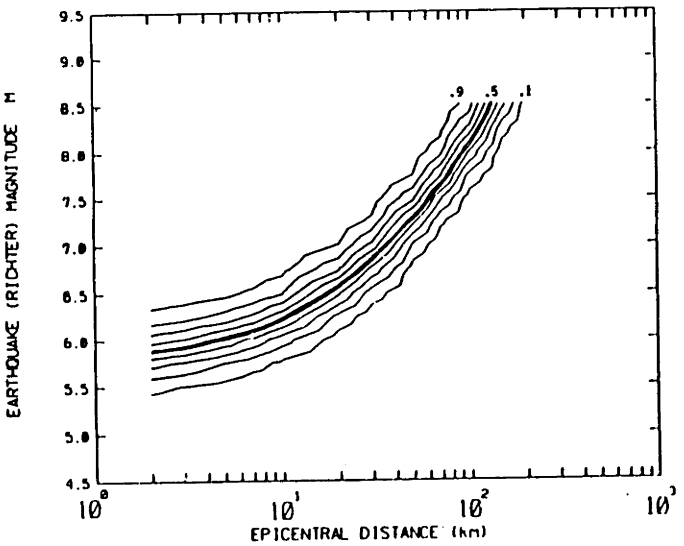


(a) No uncertainty in attenuation relationship

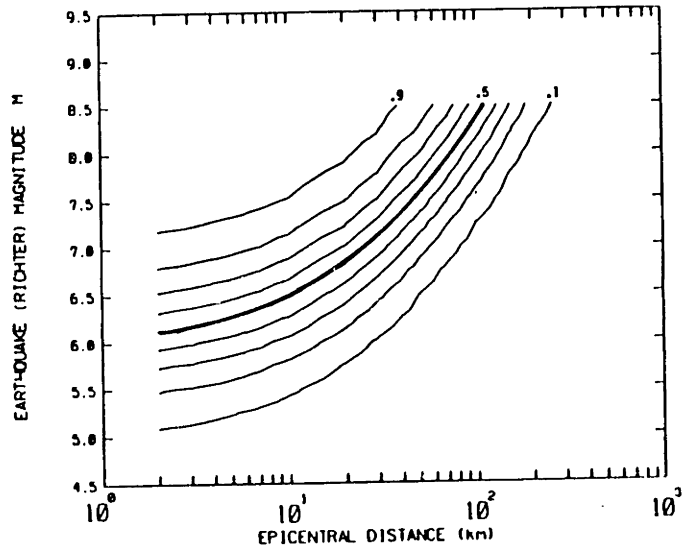


(b) Uncertainty included in attenuation relationship

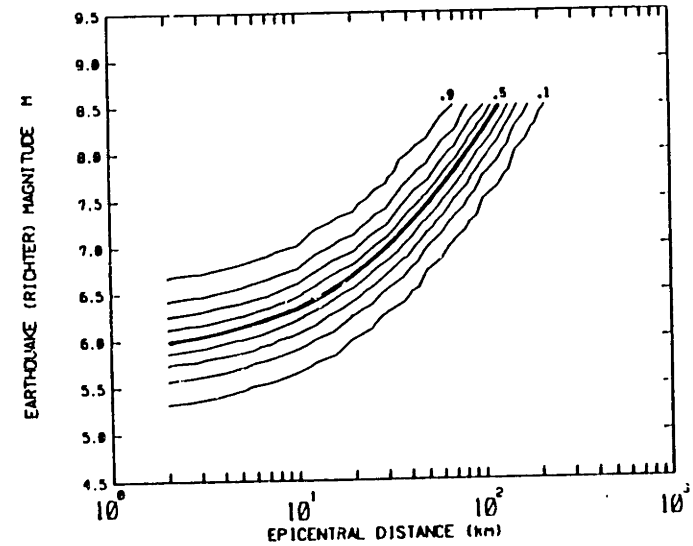
Fig. 7.8 Comparison of P = 0.1, 0.5, and 0.9 contours for clean and silty sand sites with and without uncertainty in attenuation.



(a) Clean sand

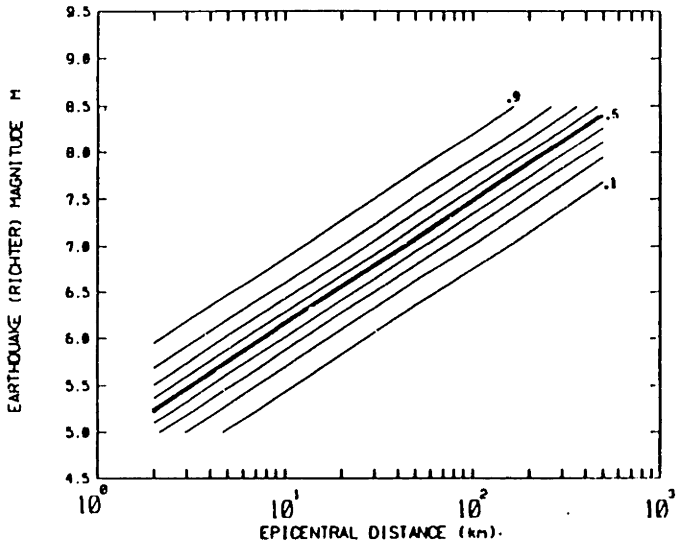


(b) Silty sand

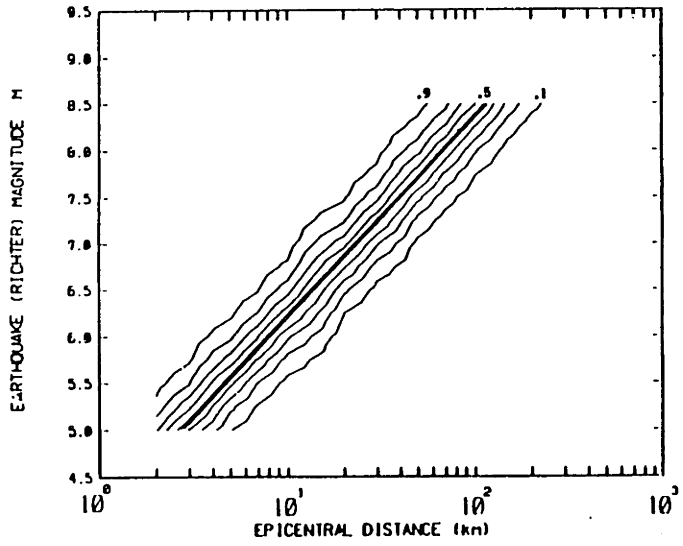


(c) Model with no distinction between clean and silty sand

Fig. 7.9 Lines of equal probability of liquefaction for a typical site, obtained using the Seed-Idriss model. Attenuation relationship from Kawashima et al. (1984). No uncertainty in attenuation assumed.

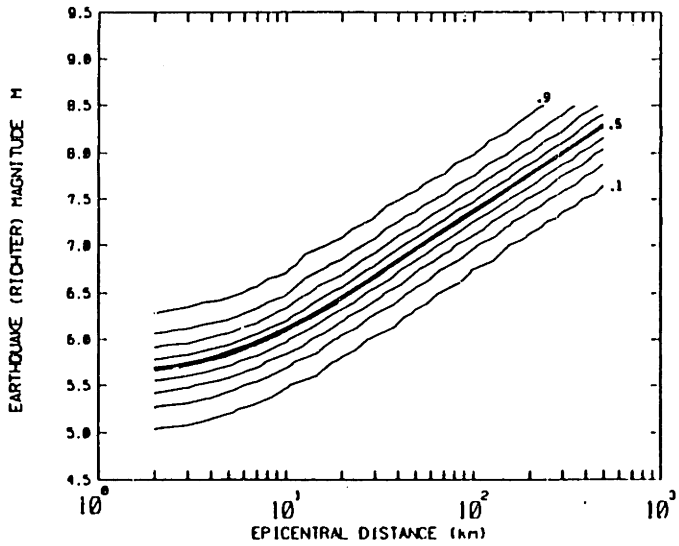


(a) Davis-Berrill epicentral model

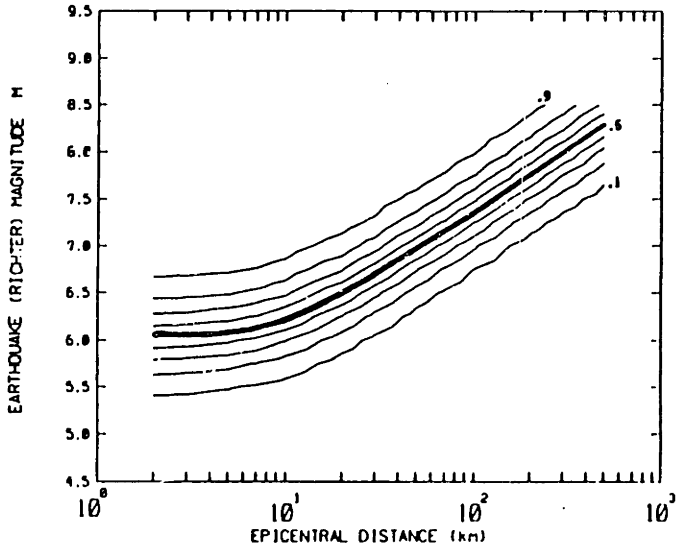


(b) Seed-Idriss base model (No distinction between clean and silty sand). Attenuation relationship from Iwasaki et al. (1978).

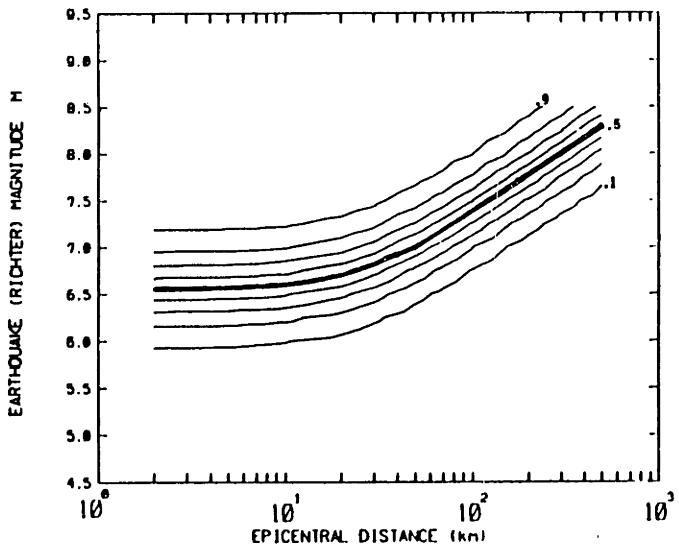
Fig. 7.10 Comparison of lines of equal probability of liquefaction using two different logistic models. Calculations performed for a typical site.



(a) Focal depth H = 5 km



(b) Focal depth H = 10 km



(c) Focal depth H = 25 km

Fig. 7.11 Lines of equal probability of liquefaction for a "standard" site obtained using the modified (hypocentral) Davis-Berrill model, assuming various focal depths.

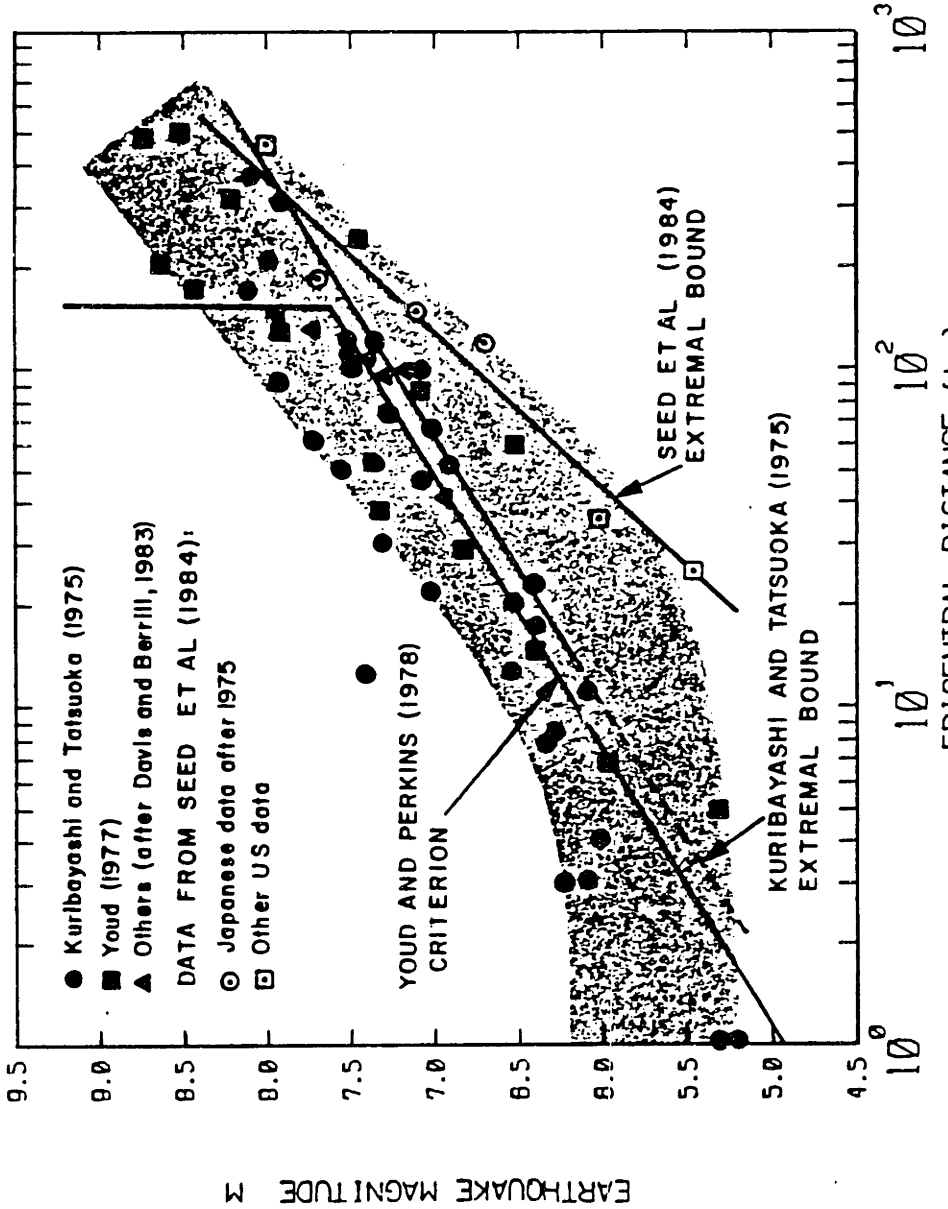


Fig. 7.12 Kuribayashi-Tatsuoka plot showing the maximum epicentral distance at which a site liquefied during an earthquake. Data added by others and various interpretations of the data are also shown.

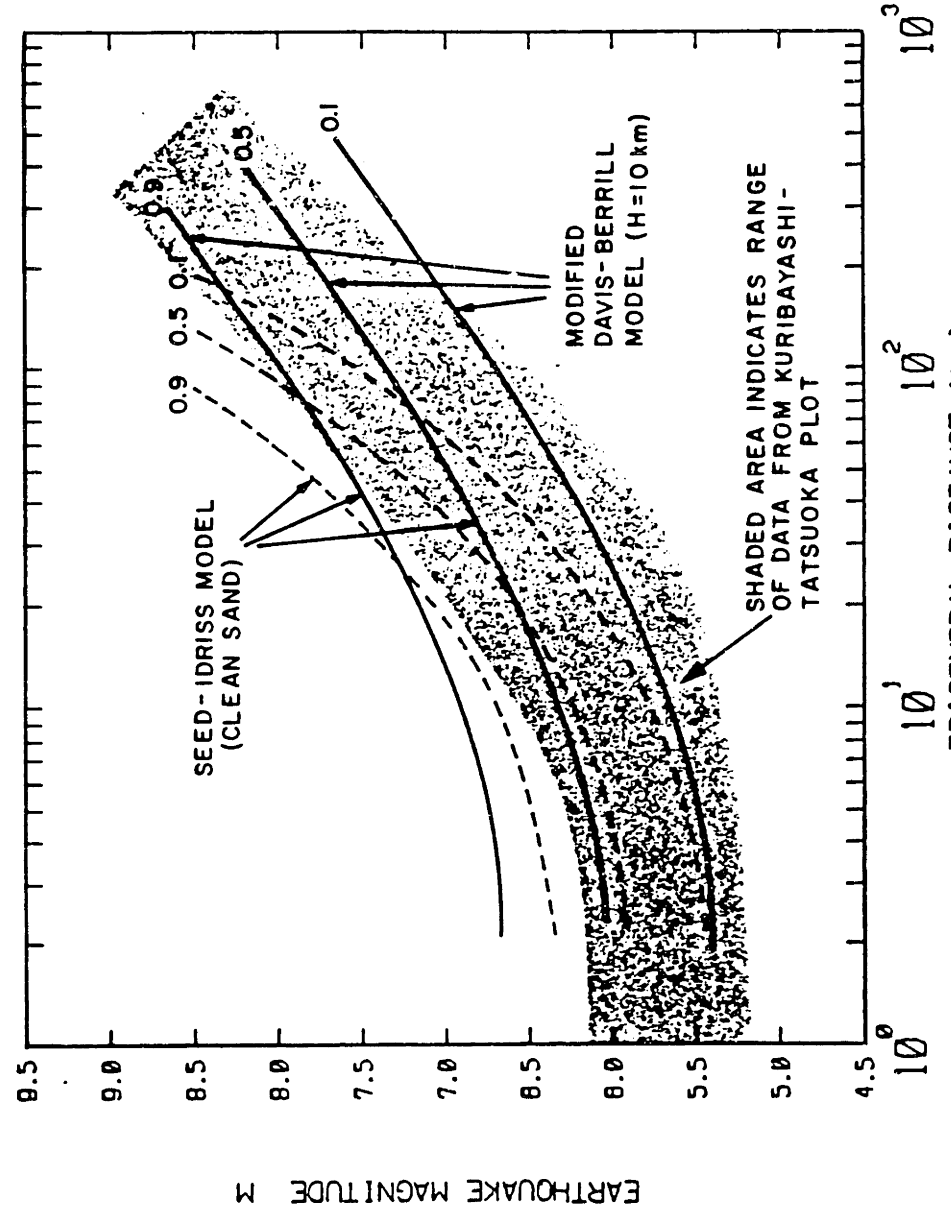


Fig. 7.13 Comparison of shapes of 0.1, 0.5, and 0.9 equi-probability contours from calculations using Seed-Idriss and modified Davis-Berrill logistic models with Kuribayashi-Tatsuoka plot.

CHAPTER 8

SUMMARY AND RECOMMENDATIONS

8.1 Summary

Statistical models have been developed to calculate the conditional probability of liquefaction as a function of earthquake load and soil resistance parameters. The models are based on the analysis of a catalog consisting of 278 cases of liquefaction and non-liquefaction occurrences during earthquakes. The data base is a synthesis of eight previous catalogs and is more comprehensive than any earlier compilation.

Binary logistic regression is the method used to obtain the models recommended for application in liquefaction risk analysis. A non-parametric kernel method was also applied to the data and used to verify the appropriateness of the assumed mathematical form of the logistic models. Modifications of the standard logistic analysis were employed to study the effects of errors and biases in the data on the fitted models. The errors and biases are due to the data collection procedures currently used in documenting liquefaction case studies.

Based on goodness-of-fit statistics and on considerations of usefulness in application, four models are recommended for risk analysis calculations. Two of them are based on the Seed and Idriss (1971) parameterization which employs the cyclic stress ratio as a measure of earthquake load. The other two use load parameters in terms of magnitude and distance, similar to that originally proposed by Davis and Berrill

(1981). Each of the four models has a domain of applicability, depending on considerations of regional seismicity and other factors, as discussed in Chapter 7. All of the models use the corrected/normalized SPT $(N_1)_{60}$ value as the measure of soil liquefaction resistance.

Several "secondary" variables that have an influence on liquefaction occurrence were identified by logistic regression analysis. The most important of these are the following soil gradation variables:

- Fines Content (FC): "Silty" sands with FC > 12% are on the average less susceptible to liquefaction, but the prediction of liquefaction or no liquefaction is less certain than for "clean" sands (FC $\leq$ 12%).
- Gravel Content (GC): Soils with GC > 10% are more susceptible to liquefaction than soils with GC $\leq$ 10% for the same SPT resistance.
- Median Grain Size (D_{50}): Increasing D_{50} causes an increase in liquefaction susceptibility for soils with the same SPT resistance.

In most regressions involving secondary variables, the accuracy of parameter estimation is limited by the small data set, so that in some cases only qualitative trends could be discerned. Of the above, only the fines content effects were considered to be adequately supported by the data to warrant its current use in risk analysis. It should be noted that the above gradation effects are most evident in models that use the Seed-Idriss parameterization. Formulations based on the Davis-Berrill load parameter (a function of magnitude and distance) are insensitive to soil gradation.

Calculations performed in Chapter 7 for a typical site using the

Seed-Idriss logistic model indicate that the difference between the probabilities of liquefaction for clean and silty sand tend to be obscured by uncertainties in the attenuation relationship. Thus, in applications such as regional mapping of liquefaction risk, obtaining such information on soil gradation may not be worthwhile.

Statistical measurement errors and biases in the data generally decrease the predictive ability of the logistic models. A significant bias, due to the propensity to investigate liquefaction cases more frequently than non-liquefaction cases, causes the logistic model to be conservative in estimating liquefaction probability. Methods are available to correct the logistic models for the effects of data errors and biases. However, definitive corrections cannot be made until additional data is obtained. To acquire such data, a practical post-earthquake sampling plan is proposed in which mapping of liquefied and non-liquefied areas is done prior to drilling borings to obtain the SPT resistance. The allocation of borings to liquefied and non-liquefied areas should then proceed as usual. The important feature of this sampling plan is that quantification of biases in the data can be derived from the method of boring allocation. More details of this sampling plan are discussed in Chapter 6.

In application, the Seed-Idriss logistic models are expected to give results consistent with deterministic methods using the Seed et al. (1984) criteria. Results of calculations using logistic models with the Davis-Berrill parameterization are, however, expected to be different from analyses using the most recent Davis and Berrill (1985) procedure. This is due to the way in which discriminant analysis is used in their

analysis. Simplistic liquefaction criteria such as that proposed by Youd and Perkins (1978) for liquefaction hazard mapping can be derived more rationally from considering statistical models of the type derived here.

Some developments of statistical methodology in this study are also worthy of note. Though methods of nonparametric kernel regression are well known, the way in which these methods are used in the analysis of liquefaction data is new (see Chapter 5). The procedures include the analysis of binary responses using a hierarchy of uniform-size and variable-size kernels, the development of a kernel method that insures monotonicity of the estimates, and optimization of kernel parameters through cross-validation. Also, some of the methods of Chapter 6 to explore the effects of errors and biases in the data are extensions of standard logistic regression.

8.2 Future Research

Refinements of the statistical models presented in this study are possible, and have been discussed within the context of the results in previous chapters. These include:

1) Site Characterization: This is the problem of how to select an SPT N_1-value that is representative of the soil deposit (see Chapter 3). The procedure recommended in this study is to use each boring as the unit of analysis and to choose the minimum N_1 as the representative value. Improvements to this procedure are certainly possible. Another aspect of site characterization relates to the existence of data dependence, e.g. among several borings located in close proximity to each other. Improved

modelling of such dependencies may be useful in determining, within a soil deposit, the physical extent of liquefaction, if it should occur.

2) Other Factors Affecting Liquefaction: Several secondary variables were identified in Chapter 4 as being statistically significant, but were not included in the proposed models due to limitations in the data. One particular factor that has not been considered in this study is the effect of the soil profile and resulting drainage conditions that may lead to dissipation of excess pore pressure. Surface manifestations of liquefaction may not be observed in some cases due to the presence of a thick non-liquefiable layer above the liquefiable stratum (Ishihara, 1985). Or, a thin layer of sand may be sandwiched between impermeable strata, and this may hinder the pore pressure dissipation. Efforts should be made to re-examine these variables as more case studies become available.

3) Corrections to Account for Data Errors and Biases: Significant adjustments to the present logistic models may result if more information is collected on errors and biases in the data. See Chapter 6 for details.

4) Treatment of Liquefaction as a Multinomial Response: This study has treated liquefaction occurrence as a binary (yes/no) response. However, one might wish to classify liquefaction according to its severity or its consequences, e.g. as suggested by Tokimatsu and Yoshimi (1983). Algorithms for multinomial logistic regression exist, which can give the probability of

occurrence of the various levels of liquefaction severity. The main difficulty in using these algorithms is the specification of the logistic model. Also, more data may be required to obtain accurate results.

5) Use of Other In-Situ Tests: In this study, the soil liquefaction resistance was measured by the standard penetration test. Though this test is relatively crude, it is one of the few measurements that provide a direct link between actual observations of liquefaction and soil properties (Peck, 1979). However, measurements from other in-situ devices, such as the cone penetrometer (CPT) or the electrical conductivity probe described by Arumoli, Arulanandan, and Seed (1985), are now being correlated to liquefaction performance. Logistic regression can also be applied to these measurements, once a significant data base has been established.

The statistical methods (especially logistic regression) employed in this study are very flexible and can be applied to many other problems in geotechnical engineering (see Carpenter, 1984; Honjo, 1985; Einstein et al., 1985). An example of a possible future application is in calculating the probability of slope failure using the factor of safety, the mode of failure, the method of measuring soil strength, etc. as explanatory variables. This is an alternative to the probabilistic methods that are currently employed.

REFERENCES

Abramowitz, M. and Stegun, I.A. (1972), Handbook of Mathematical Functions, Dover Publications, Inc., New York, Eighth Printing, 1972.

Abramson, I.S. (1982), "On Bandwidth Variation in Kernel Estimates -A Square Root Law," Annals of Statistics, Vol. 10, 1982, pp. 1217-1223.

Aitchison, J. and Aitken, G.G. (1976), "Multivariate Binary Discrimination by the Kernel Method," Biometrika, Vol. 63, No. 3, 1976, pp. 413-420.

Aitken, C.G.G. and MacDonald, D.G. (1979), "An Application of Discrete Kernel Methods to Forensic Odontology," Applied Statistics, Vol. 28, 1979, pp. 55-61.

Aki, K. and Richards, P.G. (1980), Quantitative Seismology - Theory and Methods, Vol. I, W.H. Freeman and Co., San Francisco, 1980.

Anagnos, T. and Kiremidjian, A.S (1984), "Stochastic Time - Predictable Model for Earthquake Occurrences," Bulletin of the Seismological Society of America, Vol. 74, No. 6, Dec. 1984, pp. 2593-2611.

Arulmoli, K., Arulanandan, K. and Seed, H.B. (1985), "New Method for Evaluating Liquefaction Potential," Journal of Geot. Eng., ASCE, Vol. 111, No. 1, Jan. 1985, pp. 95-114.

Atkinson, G.M., Finn, W.D.L., Charlwood, R.G. (1984), "Simple Computation of Liquefaction Probability for Seismic Hazard Applications," Earthquake Spectra, Vol. 1, No. 1, Nov. 1984, pp. 107-123.

Ben-Akiva, M. and Lerman, S. (1985), Discrete Choice Analysis: Theory and Applications to Travel Demand, The MIT Press, Cambridge, Mass., 1985.

Berkson, J. (1944), "Application of the Logistic Function to Bio-Assay," Journal of the American Statistical Assoc., Vol. 39, 1944, pp. 357-365.

Bierschwale, J.G. and Stokoe, K.H. II (1984), "Analytical Evaluation of Liquefaction Potential of Sands Subjected to the 1981 Westmorland Earthquake," Geotechnical Engineering Report GR84-15, Geotechnical Engineering Center, The University of Texas at Austin, Sept. 1984.

Bliss, C.I. (1934a), "The Method of Probits," Science, Vol. 79, 1934, pp. 38-39.

Bliss, C.I. (1934b), "The Method of Probits - A Correction," Science, Vol. 79, 1934, pp. 409-410.

Bolt, B.A. (1978), Earthquakes - A Primer, W.H. Freeman & Co., San Francisco, 1978.

Brand, R.J., Pinnock, D.E., and Jackson, K.L. (1973), "Large Sample Confidence Bands for the Logistic Response Curve and its Inverse," The American Statistician, Vol. 27, No. 4, Oct. 1973, pp. 157-160.

Breiman, L. (1985), Personal Communication to S. Liao, July 1985.

Breiman, L., Meisel, W. and Purcell, E. (1977), "Variable Kernel Estimates of Multivariate Densities," Technometrics, Vol. 19, No. 2, May 1977, pp. 135-144.

Carpenter, J.H. (1984), "Landslide Risk Along Lake Roosevelt," M.S. Thesis, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass., 1984.

Carroll, R.J., Spiegelman, C.H., Gordon, K.K., Bailey, K.T. and Abbott, R.D. (1984), "On Errors-in-Variables for Binary Regression Models," Biometrika, Vol. 71, No. 1, 1984, pp. 19-25.

Casagrande, A. (1936), "Characteristics of Cohesionless Soils Affecting the Stability of Slopes and Earth Fills," Journal of the Boston Society of Civil Engineers, Jan. 1936, Reprinted in Contributions to Soil Mechanics, 1925 to 1940, Boston Society of Civil Engineers, Oct. 1940, pp. 257-276.

Casagrande, A. (1938), "The Shearing Resistance of Soils and its Relation to the Stability of Earth Dams," Proceedings, Soils and Foundation Conference of the U.S. Engineer Department, 1938.

Casagrande, A. (1965), "Role of the 'Calculated Risk' in Earthwork and Foundation Engineering," Journal of the Soil Mechanics Division, ASCE, Vol. 91, No. SM4, July 1965, pp. 1-40.

Castro, G. (1969), "Liquefaction of Sands," Ph.D. Thesis, Harvard University, Cambridge, Mass., 1969.

Castro, G. (1975), "Liquefaction and Cyclic Mobility of Saturated Sands," Journal of the Geotechnical Engineering Division, ASCE, Vol. 101, No. GT6, June 1975, pp. 551-569.

Chameau, J.L. and Clough, G.W. (1983), "Probabilistic Pore Pressure Analysis for Seismic Loading," Journal of Geotechnical Engineering, ASCE, Vol. 109, No. 4, April 1983, pp. 507-524.

Chow, Y.S., Geman, S. and Wu, L.D. (1983), "Consistent Cross-Validated Density Estimation," Annals of Statistics, 1983, Vol. 11, pp. 25-38.

Christian, J.T. and Swiger, W.F. (1975), "Statistics of Liquefaction and SPT Results," Journal of Geotechnical Engineering Div., ASCE, Vol. 101, No. GT11, Nov. 1975, pp. 1135-1150.

Cornell, C.A. (1968), "Engineering Seismic Risk Analysis," Bulletin of the Seismological Society of America, Vol. 58, No. 5, Oct. 1968, pp. 1583-1606.

Cornell, C.A. (1971), "Probabilistic Analysis of Damage to Structures Under Seismic Loads," Dynamic Waves in Civil Engineering, D.A. Howells, I.P. Haigh and C. Taylor, Eds., Wiley-Interscience, London, 1971.

Cox, D.R. (1970), The Analysis of Binary Data, Methuen & Co. Ltd., London, 1970; Reprinted by Chapman and Hall, London, 1983.

Daganzo, C. (1979), Multinomial Probit - The Theory and its Application to Demand Forecasting, Academic Press, New York, 1979.

Davis, R.O. and Berrill, J.B. (1981), "Assessment of Liquefaction Potential Based on Seismic Energy Dissipation," Proc., Int'l. Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 1, pp. 187-190.

Davis, R.O. and Berrill, J.B. (1982), "Energy Dissipation and Seismic Liquefaction in Sands," Earthquake Engineering and Structural Dynamics, 1982, Vol. 10, pp. 59-68.

Davis, R.O. and Berrill, J.B. (1983), "Comparison of a Liquefaction Theory with Field Observations," Geotechnique, Vol. 33, No. 4, Dec. 1983, pp. 455-460.

Der Kiureghian, A. and Ang, A. H-S. (1977), "Fault Rupture Model for Seismic Risk Analysis," Bulletin of the Seismological Society of America, Vol. 67, No. 4, Aug. 1977, pp. 1173-1194.

Dobry, R., Ladd, R.S., Yokel, F.Y., Chung, R.M. and Powell, D. (1982), "Prediction of Pore Water Pressure Buildup and Liquefaction of Sands during Earthquakes by the Cyclic Strain Method," NBS Building Science Series 138, National Bureau of Standards, U.S. Dept. of Commerce, U.S. Govt. Printing Office, Washington, D.C., July 1982.

Dobson, A.J. (1983), An Introduction to Statistical Modelling, Chapman and Hall, London, 1983.

Donovan, N.C. (1971), "A Stochastic Approach to the Seismic Liquefaction Problem," Proc., Int'l. Conference on Applications of Statistics and Probability to Soil and Structural Engineering, Hong Kong, 1971, pp. 513-535.

EERI (1980), "Reconnaissance Report, Montenegro, Yugoslavia Earthquake, April 15, 1979," Earthquake Engineering Research Institute, Nov. 1980.

Easterling, R.G. and Heller, L.W. (1976), "Discussion - Statistics of Liquefaction and SPT Results," Journal of the Geotechnical Engineering Div., ASCE, Vol. 102, No. GT10, Oct. 1976, pp. 1126-1128.

Einstein, H.H., Baecher, G.B. and Veneziano, D. (1985), "Statistical Correlation Analysis of Downhole Measurement Data," Draft report submitted to Nationale Genossenschaft für die Lagerung Radioaktiver Abfälle (NAGRA), May 1985.

Efron, B. (1982), The Jacknife, The Bootstrap and Other Resampling Plans,

Society of Industrial and Applied Mathematics, Philadelphia, PA, 1982.

Efron, B. and Gong, G. (1983), "A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation," The American Statistician, Vol. 37, No. 1, Feb. 1983, pp. 36-48.

Epanechnikov, V.A. (1969), "Nonparametric Estimates of a Multivariate Probability Density," Theory of Probability and its Applications, Vol. 14, 1969, pp. 153-158.

Fardis, M.N. and Veneziano, D. (1982), "Probabilistic Analysis of Deposit Liquefaction," Journal of the Geotechnical Engineering Div., ASCE, Vol. 108, No. GT3, March 1982, pp. 395-417.

Finney, D.J. (1971), Probit Analysis, 3rd edition, Cambridge University Press, 1971.

Fuller, W.A. (1980), "Properties of Some Estimators for Errors-in-Variables Model," Annals of Statistics, Vol. 8, 1980, pp. 407-422.

Ganse, R.A., Amemiya, Y. and Fuller, W.A. (1983), "Prediction When Both Variables are Subject to Error, With Application to Earthquake Magnitudes," Journal of the American Statistical Assoc., Vol. 78, No. 384, Dec. 1983, pp. 761-765.

Gleser, L.R. (1981), "Estimation in a Multivariate 'Errors-in-Variables' Regression Model: Large Sample Results," Annals of Statistics, Vol. 9, 1981, pp. 24-44.

Good, I.J. and Gaskins, R.A. (1971), "Nonparametric Roughness Penalties for Probability Densities," Biometrika, Vol. 58, 1971, pp. 255-277.

Gu, W. and Wang, Y. (1984), "An Approach to the Quadratic Nonlinear Formulae for Predicting Earthquake Liquefaction Potential by Stepwise Discriminant Analysis," Proc., 8th World Conference on Earthquake Engineering, San Francisco, July 1984, Vol. 3, pp. 119-126.

Habbema, J.D.F., Hermans, J. and van Den Broek, K. (1974), "A Stepwise Discriminant Analysis Program using Density Estimation," Compstat 1974, ed. by G. Bruckmann, Physica Verlag, Vienna, 1974, pp. 101-110.

Haldar, A. and Tang, W.H. (1979), "Probabilistic Evaluation of Liquefaction Potential," Journal of the Geotechnical Engineering Div., ASCE, Vol. 105, No. GT2, Feb. 1979, pp. 145-163.

Hand, D.J. (1981), Discrimination and Classification, John Wiley & Sons, 1981.

Hand, D.J. (1982), Kernel Discriminant Analysis, John Wiley & Sons, 1982.

Hanushek, E.A. and Jackson, J.E. (1977), Statistical Methods for Social Scientists, Academic Press, New York, NY, 1977, 374 p.

Härdle, W. and Marron, J.S. (1985), "Optimal Bandwidth Selection in Nonparametric Regression," Annals of Statistics, Vol. 13, No. 4, Dec. 1985, pp. 1465-1481.

Hastie, T.J. (1983), "Non-parametric Logistic Regression," Report, Dept. of Statistics, Stanford University, Stanford, CA, 1983.

Hauck, W.W. (1983), "A Note on Confidence Bounds for the Logistic Response Curve," The American Statistician, Vol. 37, No. 2, May 1983, pp. 158-160.

Heidari, M. (1981), "Centrifugal Modelling of Earthquake Induced Liquefaction in Saturated Sand," M. Phil. Thesis. Cambridge University, England, 1981.

Heidari, M. (1986), Ph.D. Thesis in Progress, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass., 1986.

Hensher, D.A. and Johnson, L.W. (1981), Applied Discrete-Choice Modelling, Halsted Press/John Wiley & Sons, New York, 1981.

Honjo, Y. (1985), "Dam Filters: Physical Behavior, Probability of Malfunctioning, and Design Criteria," Ph.D. Thesis, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass., 1985.

Hoose, S.N., Wilson, R.C., Rosenfeld, J.H., "Liquefaction-Caused Ground Failure During the February 4, 1976, Guatemala Earthquake," Proc. of the Intl. Symposium on the February 4, 1976, Guatemalan Earthquake and the Reconstruction Process, Vol. 2, , 1978.

Horowitz, J.L. (1982), "Evaluation of Usefulness of Two Standard Goodness-of-Fit Indicators for Comparing Non-Nested Random Utility Models," Advances in Trip Generation, Transportation Research Record 874, 1982, pp. 19-25.

Idriss, I.M. (1978), "Characteristics of Earthquake Ground Motions," Proc. ASCE Specialty Conference on Earthquake Engineering and Soil Dynamics, Pasadena, Calif., 1978, Vol. 3, pp. 1151-1265.

Ishihara, K. (1985), "Stability of Natural Deposits during Earthquakes," Proc., XI International Conference on Soil Mechanics and Foundation Engineering, San Francisco, 1985, Vol. 1, pp. 321-376.

Iwasaki, T., Tatsuoka, F., Tokida, K. and Yasuda, S. (1978), "A Practical Method for Assessing Soil Liquefaction Potential," Proc., 2nd Intl. Conference on Microzonation for Safer Construction - Research and Application, San Francisco, 1978, Vol. 2, pp. 885-896.

Jaime, A., Montanez, L. and Romo, M.P. (1981), "Liquefaction of Enmedio Island Soil Deposits," Proc., Intl. Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 1, pp. 529-534.

Johnson, N.L. and Kotz, S. (1970), Distributions in Statistics: Continuous Univariate Distributions - 2, Vol. 3, Houghton Mifflin Co., Boston, 1970.

Johnson, R.A. and Wichern, D.W. (1982), Applied Multivariate Statistical Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1982.

Joyner, W.B. and Boore, D.M. (1981), "Peak Horizontal Acceleration and Velocity from Strong-Motion Records Including Records from the 1979 Imperial Valley, California, Earthquake," Bulletin of the Seismological Society of America, Vol. 71, No. 6, Dec. 1981, pp. 2011-2038.

Kanamori, H. and Jennings, P.C. (1978), "Determination of Local Magnitude, M_L, from Strong-Motion Accelerograms," Bulletin of the Seismological Society of America, Vol. 68, No. 2, April 1978, pp. 471-485.

Kaufman, L.P. (1981), "Percentage Silt Content in Sands and its Effect on Liquefaction Potential," M.S. Thesis, Dept. of Civil and Urban Engineering, Univ. of Colorado, Boulder, CO; also published as a Technical Report by the Bureau of Reclamation, Denver, Colorado, 1981.

Kavazanjian, E., Roth, R.A. and Echezuria, H. (1985), "Liquefaction Potential Mapping for San Francisco," Journal of the Geotechnical Engineering Div., ASCE, Vol. 111, No. 1, Jan. 1985, pp. 54-76.

Kawashima, K., Aizawa, K. and Takahashi, K. (1984), "Attenuation of Peak Ground Motion and Absolute Acceleration Response Spectra," Proc., 8th World Conference on Earthquake Engineering, San Francisco, July 1984, Vol. 2, pp. 257-264.

Kiremidjian, A.S. and Anagnos, T. (1984), "Stochastic Slip-Predictable Model for Earthquake Occurrences," Bulletin of the Seismological Society of America, Vol. 74, No. 2, April 1984, pp. 739-755.

Koizumi, Y. (1966), "Changes in Density of Sand Subsoil Caused by the Niigata Earthquake," Soils and Foundations, Vol. 6, No. 2, 1966, pp. 38-44.

Kuribayashi, E. and Tatsuoka, T. (1975), "Brief Review of Liquefaction During Earthquakes in Japan," Soils and Foundations, Vol. 15, No. 4, Dec. 1975, pp. 81-92.

Lauder, L.J. (1983), "Direct Kernel Assessment of Diagnostic Probabilities," Biometrika, Vol. 70, No. 1, 1983, pp. 251-256.

Lee, K.L. and Fitton, J.A. (1969), "Factors Affecting the Cyclic Loading Strength of Soil," Vibration Effects of Earthquakes on Soils and Foundations, ASTM STP 450, American Society for Testing and Materials, 1969, pp. 71-95.

Lerman, S. (1983), Personal Communication with S. Liao, 1983.

Liao, S. and Whitman, R.V. (1986), "Overburden Correction Factors for SPT in Sand," to be published in the Journal of Geotechnical Engineering, ASCE, March 1986.

Madansky, A. (1959), "The Fitting of Straight Lines when Both Variables are Subject to Error," Journal of the American Statistical Association, Vol. 54, 1959, pp. 173-205.

Manski, C.F. and Lerman, S.R. (1977), "The Estimation of Choice Probabilities from Choice-Based Samples," Econometrica, Vol. 45, No. 8, Nov. 1977, pp. 1977-1988.

Marron, J.S. (1985), "An Asymptotically Efficient Solution to the Bandwidth Problem," Annals of Statistics, Vol. 13, No. 3, Sept. 1985, pp. 1011-1023.

McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Econometrics, ed. by P. Zarembka, Academic Press, New York, 1974, pp. 105-142.

McFadden, D. (1976a), "Quantal Choice Analysis: A Survey," Annals of Economic and Social Measurement, Vol. 5, No. 4, 1976, pp. 363-390.

McFadden, D. (1976b), "A Comment on Discriminant Analysis 'Versus' Logit Analysis," Annals of Economic and Social Measurement, Vol. 5, No. 4, 1976, pp. 511-523.

McGuire, R.K., Tatsuoka, F., Iwasaki, F. and Tokida, K. (1978), "Probabilistic Procedures for Assessing Soil Liquefaction Potential," Journal of Research, Public Works Research Institute, Tokyo, Japan, Vol. 19, 1978, pp. 1-38.

McGuire, R.K., Tatsuoka, F., Iwasaki, F. and Tokida, K. (1979), "Assessment of Probability of Liquefaction of Water-Saturated Reclaimed Land," Proc., 3rd Intl. Conf. on Applications of Statistics and Probability in Soil and Structural Engineering, Sydney, Australia, 1979, pp. 786-801.

Michalek, J.E. and Tripathi, R.C. (1980), "The Effect of Errors of Diagnosis and Measurement on the Estimation of the Probability of an Event," Journal of the American Statistical Association, Vol. 75, 1980, pp. 713-721.

NRC (1985), Liquefaction of Soils During Earthquakes, Report of a Workshop Held at M.I.T. Endicott House, Dedham, MA, March, 1985, Published by the National Research Council, Committee on Earthquake Engineering, National Academy Press, Washington, D.C., 1985.

Nadaraya, E.A. (1964), "On Estimating Regression," Theory of Probability and its Applications, Vol. 9, 1964, pp. 141-142.

Patwardan, A.S., Kulkarni, R.B. and Tocher, D. (1980), "A Semi-Markov Model for Characterizing Recurrence of Great Earthquakes," Bulletin of the Seismological Society of America, Vol. 70, 1980, pp. 323-347.

Pearl, R. and Reed, L.J. (1920), "On the Rate of Growth of the Population of the United States since 1790 and its Mathematical Representation," Proceedings of National Academy of Sciences, Vol. 6, 1920, pp. 275-288.

Peck, R.B. (1979), "Liquefaction Potential: Science Versus Practice," Journal of the Geotechnical Engineering Div., ASCE, Vol. 105, No. GT3, March 1979, pp. 393-398.

Prentice, R.L. and Pyke, R. (1979), "Logistic Disease Incidence Models and Case-Control Studies," Biometrica, Vol. 66, No. 3, 1979, pp. 403-411.

Rice, J. (1984), "Bandwidth Choice for Nonparametric Regression," Annals of Statistics, Vol. 12, No. 4, 1984, pp. 1215-1230.

Rudemo, M. (1982), "Empirical Choice of Histograms and Kernel Density Estimators," Scandinavian Journal of Statistics - Theory and Applications, Vol. 9, No. 2, 1984, pp. 65-78.

Schwartz, D.P. and Coppersmith, K.J. (1984), "Fault Behavior and Characteristic Earthquakes: Examples from the Wasatch and San Andreas Faults," Journal of Geophysical Research, Vol. 89, No. B7, 1984, pp. 5681-5698.

Scott, D.W., Tapia, R.A. and Thompson, J.R. (1977), "Kernel Density Estimation Revisited," Nonlinear Analysis: Theory, Methods and Applications, Vol. 1, 1977, pp. 339-372.

Seed, H.B. (1979), "Soil Liquefaction and Cyclic Mobility Evaluation for Level Ground During Earthquakes," Journal of Geotechnical Engineering Div., ASCE, Vol. 105, No. GT2, Feb. 1979, pp. 201-255.

Seed, H.B., Arango, I. and Chan, C.K. (1975), "Evaluation of Soil Liquefaction Potential During Earthquakes," Report No. EERC 75-28, Earthquake Engineering Research Center, Univ. of California, Berkeley, CA, Oct. 1975.

Seed, H.B., Arango, I., Chan, C.K., Gomez-Masso, A. and Grant-Ascow, R. (1981), "Earthquake-Induced Liquefaction Near Lake Amatitlan, Guatemala, Journal of the Geotechnical Engineering Div., ASCE, Vol. 107, No. GT4, April 1981, pp. 501-518.

Seed, H.B. and Idriss, I.M. (1971), "Simplified Procedure for Evaluating Soil Liquefaction Potential," Journal of the Soil Mechanics and Foundations Div., ASCE, Vol. 97, No. SM9, Sept. 1971, pp. 1249-1273.

Seed, H.B., Idriss, I.M. and Arango, I. (1983), "Evaluation of Liquefaction Potential Using Field Performance Data," Journal of Geotechnical Engineering, ASCE, Vol. 109, No. 3, Mar. 1983, pp. 458-482.

Seed, H.B., Martin, P.P. and Lysmer, J. (1976), "Pore Pressure Changes During Soil Liquefaction" Journal of the Geotechnical Engineering Div., ASCE, Vol. 102, No. GT4, April 1976, pp. 323-346.

Seed, H.B., Tokimatsu, K., Harder, L.F. and Chung, R.M. (1984), "The Influence of SPT Procedures in Soil Liquefaction Resistance Evaluations," Report No. UBC/EERC 84-15, Earthquake Engineering Research Center, Univ. of California, Berkeley, CA, Oct. 1984.

Seed, H.B., Tokimatsu, K., Harder, L.F. and Chung, R.M. (1985), "Influence of SPT Procedures in Soil Liquefaction Resistance Evaluations," Journal of Geotechnical Engineering, ASCE, Vol. 111, No. 12, Dec. 1985, pp. 1425-1445.

Sherif, M.A., Tien, Y.B. and Pan, Y.W. (1983), "Liquefaction Potential of Silty-Sand," Soil Engineering Research Report No. 26, Dept. of Civil Engineering, Univ. of Washington, Seattle, Washington, Dec. 1983.

Shibata, T. (1981), "Relations Between N-Value and Liquefaction Potential of Sand Deposits," Proceedings, 16th Annual Convention of Japanese Society of Soil Mechanics and Foundation Engineering, 1981, pp. 621-624 (in Japanese).

Silverman, B.W. (1978), "Choosing the Window Width When Estimating a Density," Biometrika, Vol. 65, No. 1, 1978, pp. 1-11.

Silverman, B.W. (1984), "Spline Smoothing: The Equivalent Kernel Method," Annals of Statistics, Vol. 12, No. 3, 1984, pp. 898-916.

Stefanski, L.A. and Carroll, R.J. (1985), "Covariate Measurement Error in Logistic Regression," Annals of Statistics, Vol. 13, No. 4, 1985, pp. 1335-1351.

Sykes, L.R. and Nishenko, S. (1984), "Probabilities of Occurrence of Large Plate Rupturing Earthquakes for the San Andreas, Jacinto, and Imperial Faults, California," Journal of Geophysical Research, Vol. 89, 1984, pp. 5905-5927.

Tanimoto, K. (1977), "Evaluation of Liquefaction of Sandy Deposits by a Statistical Method," Proc., 6th World Conference on Earthquake Engineering, New Delhi, 1977, Vol. 3, pp. 2201-2206.

Tanimoto, K. and Noda, T. (1976), "Prediction of Liquefaction Occurrence of Sandy Deposits During Earthquake by a Statistical Method," Proc., Japan Society of Civil Engineers, No. 256, Dec. 1976.

Tapia, R.A. and Thompson, J.R. (1978), Nonparametric Probability Density Estimation, Johns Hopkins University Press, Baltimore, MD, 1978.

Tatsuoka, F., Iwasaki, T. et al. (1980), "Standard Penetration Tests and Soil Liquefaction Potential Evaluation," Soils and Foundations, Vol. 20, No. 4, Dec. 1980, pp. 95-111.

Theil, H. (1969), "A Multinomial Extension of the Linear Logit Model," International Economic Review, Vol. 10, 1969, pp. 251-259.

Theil, H. (1971), Principles of Econometrics, John Wiley and Sons, Inc., New York, 1971.

Tokimatsu, K. and Yoshimi, Y. (1981), "Field Correlation of Soil Liquefaction with SPT and Grain Size," Proceedings, Intl. Conf. on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 1, pp. 203-208.

Tokimatsu, K. and Yoshimi, Y. (1983), "Empirical Correlation of Soil Liquefaction Based on SPT N-Value and Fines Content," Soils and Foundations, Vol. 23, No. 4, Dec. 1983, pp. 56-74.

Veneziano, D. and Liao, S. (1984), "Statistical Analysis of Liquefaction Data," Proceedings, 4th ASCE Specialty Conf. on Probabilistic Mechanics and Structural Reliability, Berkeley, CA, Jan. 1984, pp. 206-209.

Walker, S.H. and Duncan, D.B. (1967), "Estimation of the Probability of an Event as a Function of Several Independent Variables," Biometrica, Vol. 54, 1967, pp. 167-179.

Wang, Y., Luan, F., Han, Q. and Li, G. (1980), "Formulae for Predicting Liquefaction Potential of Clayey Silt as Derived from a Statistical Method," Proceedings, 7th World Conf. on Earthquake Engineering, Istanbul, 1980, Vol. 3, pp. 227-234.

Watson, G.S. (1964), "Smooth Regression Analysis," Sankhyā, Series A, Vol. 26, pp. 359-372.

Whitman, R.V. (1971), "Resistance of Soil to Liquefaction and Settlement," Soils and Foundations, Vol. 11, No. 4, Dec. 1971, pp. 59-68.

Whitman, R.V. (1984), "Evaluating Calculated Risk in Geotechnical Engineering," Journal of Geotechnical Engineering, ASCE, Vol. 110, No. 2, Feb. 1984, pp. 145-188.

Whitman, R.V. (1985), "On Liquefaction," Proceedings, XIth Intl. Conf. on Soil Mechanics and Foundation Engineering, San Francisco, 1985, Vol. 4, pp. 1923-1926.

Wong, W.H. (1983), "On the Consistency of Cross-Validation in Kernel Nonparametric Regression," Annals of Statistics, 1983, Vol. 11, pp. 1136-1141.

Xie, J. (1979), "Empirical Criteria of Sand Liquefaction," The 1976 Tangshan China Earthquake---Papers Presented at the 2nd. U.S. National Conf. on Earthquake Engineering, Stanford University, 1979, published by Earthquake Engineering Research Institute, Berkeley, CA, Stanford University, March 1980, pp. 89-101.

Yegian, M. (1976), "Risk Analysis for Earthquake-Induced Ground Failure by Liquefaction," Ph.D. Thesis, Massachusetts Institute of Technology, Dept. of Civil Engineering, 1976.

Yegian, M.K. and Vitelli, B.M. (1981a), "Probabilistic Analysis for Liquefaction," Report No. CE-81-1, Northeastern University, Boston, Mass., March 1981.

Yegian, M.K. and Vitelli, B.M. (1981b), "Analysis for Liquefaction: Empirical Approach," Proc., Intl. Conf. on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 1, pp. 173-177.

Yegian, M. and Whitman, R.V. (1978), "Risk Analysis for Ground Failure by Liquefaction," Journal of the Geotechnical Engineering Div., ASCE, Vol. 104, No. GT7, July 1979, pp. 921-938.

Yoshimi, Y., Richart, F.E., Jr., Prakash, S., Barkan, D.D. and Ilyichev, V.A. (1977), "Soil Dynamics and Its Application to Foundation Engineering," State-of-the-Art Report, IXth Int. Conference on Soil Mechanics and Foundation Engineering, Tokyo, 1977, Vol. 2, pp. 605-650.

Youd, T.L. (1977), Discussion, "Brief Review of Liquefaction During Earthquakes in Japan," by E. Kuribayashi and F. Tatsuoka, Soils and Foundations, Vol. 17, No. 1, Mar. 1977, pp. 82-85.

Youd, T.L. (1984), "Recurrence of Liquefaction at the Same Site," Proc., 8th World Conf. on Earthquake Engineering, San Francisco, Calif., July 1984, Vol. 3, pp. 231-238.

Youd, T.L. and Perkins, M. (1978), "Mapping Liquefaction-Induced Ground Failure Potential," Journal of the Geotechnical Engineering Div., ASCE, Vol. 104, No. GT4, April 1978, pp. 433-446.

APPENDIX A

LIQUEFACTION DATA CATALOG

A.1 Guide to the Catalog

The liquefaction catalog presented in this appendix consists of five tables:

Table A.1 - Case Identification and Qualitative Attributes

Table A.2 - Edited/Enhanced Data

Table A.3 - Previously Published Catalog Identification Codes

Table A.4 - Data from Previously Published Catalogs

Table A.5 - Case Source Reference Guide

Table A.6 - Magnitude Scales for Tabulated Earthquakes

Tables A.1 and A.2 are the main products of the data compilation and contain the data analyzed in this study. Tables A.3 through A.6 are provided for documentation purposes only. In the tables, each case study is identified by a four digit code. The first two digits associate the case with a particular earthquake, and the second two digits identify the particular site affected by the earthquake.

Section A.2 provides commentaries on selected case studies, and Section A.3 provides a key to the tables. Following the tables is a bibliography of the original source references that provided the data for this catalog.

A general descritpion of the general philosophies, problems, and difficulties in the compilation of this catalog is presented in Chapter 3.

A.2 Selected Case Commentaries

The commentaries in this section refer to the editing and enhancement of the data from Tables A.3 and A.4 to compile Tables A.1 and A.2. Catalog numbers mentioned in the comments numbers refer to the source catalogs (see Table 3.1). Most of the comments are very brief, and source references may have to be consulted to understand them. Furthermore, the comments are not purported to be complete in describing every nuance of the catalog. The primary reason for these commentaries is to provide a guide to future researchers who may wish to re-examine and possibly revise some of the entires in Tables A.1 and A.2.

Cases 0101 to 0102

Assume reported Kawasumi magnitude equivalent to Richter magnitude. Assume epicentral distance equal to distance to energy release given by Catalog 2. Assume gradation same as cases for Niigata (1964) given by Catalog 7.

Cases 0201 to 0202

Same comments as for Cases 0101 to 0102.

Case 0301

Revised fines content reported by Catalog 7, based on Kishida (19).

Cases 0303

Catalog 2 lists this as a case of no liquefaction, which is an error, based on Kishida (1966).

Case 0305

Excluded from catalog because of boring correlation with case 0302.

Cases 0501 to 0502

Source catalog gives focal depth of 25 km and hypocentral distance of 16 km, which is inconsistent. Use distance data from Clough and Chameau (1983).

Case 504

Use Catalog 6 data. Catalog 3 and 4 source references not available and N-value probably bsed on D&M sampler.

Case 503

Excluded from catalog because data don't make sense.

Case 504

Ground water depth variable. Gross assumption necessary.

Case 0506

Ground water reported to be higher at time of earthquake than at time of boring. To estimate accelerations at the site, used Joyner & Boore (1981) equation and a moment magnitude of 7.7 as reported by Hanks and Kanamori (1979). DER obtained from Youd and Hoose (1978).

Cases 507 to 508

Water table not reported by Clough and Chaneau (1983). Back-calculated based on available data.

Cases 701 to 706

Excluded from catalog because data from same borings as in other cases.

Case 801

Relative density used by source catalogs to estimate SPT resistance.

Case 903

This case is identified as a non-liquefaction observation by Catalog 6, but may actually be a liquefaction data point. Source reference Kuribayashi et al. (1977) not clear on this. Use non-liquefaction classification.

Cases 1001 to 1004

SPT resistance possibly based on D&M sample resistance. Acceleration data based on Pyke et al. (1978). Disregard distance data from Catalog 4; use data from Pyke et al. (1978). Compared N-values reported by Catalog 4 to boring data from D&M (1977) and found no direct correspondence.

Cases 1101 to 1103

Relative density used by source catalogs to estimate SPT resistance.

Case 1203

Catalog 4 reports this as a no liquefaction case, which is an error, based on Kishida (1969). Note N-values reported by Catalog 4 also appear to be in error.

Cases 1301 to 1303

Case 1301 and 1302 based on same boring. Note, however the possibility of error in source reference by Kishida (1969). Kishida reports Takaya 45 (1303) is in a paddy field and that Takaya 2 (1301 & 1302) is in an area with ground elevation 1 m higher than the paddy field. But water table in Takaya 45 is lower than in Takaya 2, implying Takaya 2 is actually the paddy field boring. Could the boring logs have been

inadvertently switched?. Because of this possible confusion and also divergence of interpretation of representative N-values by various catalogers (see Table A.4), it was decided to exclude these cases from the catalog.

Case 1305

Catalogs 2 and 4 report this as a liquefaction case study, which is possibly in error, based on Kishida (1969). Also, there is a divergence of interpretation of representative N-values for this case by various catalogers (see Table A.4). Because of this and because Case 1305 was found to be a severe statistical outlier during preliminary analyses, it was decided to exclude this case from the final catalog.

Case 1401

Assume epicentral distance equal to hypocentral distance reported by Catalog 4.

Case 1501-1534

Assume epicentral distance equal to hypocentral distance reported by Catalog 4. For Case 1505, assume D_{50} and U_C same as in Cases 1535 and 1536.

Case 1706

Excluded because case is the same as 1701 and/or 1702, even though no common site or case code is assigned in Table A.

Case 1801 to 1823

The cases tabulated for the Niigata (1964) earthquake overlap with each other, i.e., several are based on the same set of borings, but it is not

clear which are correlated. Decided to base enhanced data largely on Catalog 7 because the data grouping in that catalog seemed to preserve data independence well; at least geographically, the boring groups were distinctly separated.

Case 2001

Assumed fines content equal to 5%, based on classification of this case as an SP soil in Catalog 8.

Cases 2101 to 2102

Cases excluded from final catalog because the sand ("shirasu") has unusual characteristics, is not a silica sand, and SPT resistance values based on correlation with unusual sounding technique (not SPT).

Cases 2201 to 2213

The cases tabluated for the Tokachi-Oki Earthquake overlap with each other, i.e., several are based on the same set of borings. Sorted out and excluded repeated/correlated data. Re-evaluated epicentral distance and distance to energy release, based on an "average" value of conflicting data from source reference. For Case 2213, it appears that the boring shows data prior to site improvement, which is not representative of the site at the time of the earthquake.

Cases 2301 to 2306

Back-calculated epicentral distance from hypocentral distance and focal depth given by Catalog 4. Assumed distance is energy release equal to epicentral distance.

Case 2501

Acceleration data based on response analysis reported in Teczan et al. (1977).

Cases 2604 to 2611

Cases excluded because some data overlap with Cases 2601 to 2603. Also, many data considered unreliable. Note that Yegian (Catalog 3) reports these cases, but Yegian and Vitelli (Catalog 4) chose to exclude many of them.

Case 2701

Two borings reported by source reference Sato et al. (1973), but they were performed prior to construction of pile foundation for storage tank. Use site average N-value, assuming this represents not soil under the tank, but adjacent to it.

Case 3101 to 3102

Case 3102 found to be a statistical outlier during preliminary analysis. Also, 3101 and 3102 N-values based on an unusual sounding method (not SPT). Decided to exclude both cases from final catalog.

Cases 3003 to 3005

From source reference (Zhou, 1981), appears that the average values obtained are averaged over many sites, and also are correlations with cone data. Excluded from final catalog because data does not conform to catalog format.

Case 3008

Found to be a severe outlier during preliminary statistical analyses. Excluded this case from final catalog.

Cases 3401 to 3425

Used same data as originally tabulated for larger Miyagiken-Oki event, noting that this first event of smaller magnitude caused no liquefaction.

Case 3801

Used earthquake data from EERI (1980). Note that distance to energy release assumed equal to epicentral distance because fault zone thought to be perpendicular to coast. Use layer average because N-value decrease consistently with depth, indicating possible soil grain size variation.

Case 4001

Data from Diaz-Rodriguez (1984) seems to show that Town of Delta is right on the fault rupture, but this is not exactly clear from the paper. Used average acceleration values based on the Imperial Valley 1979 Earthquake.

Case 4102

Re-evaluated percent fines data reported in Catalog 7, based on Ishihara (1981).

A.3 Key to Catalog Tables

This section describes the variables, notations, and abbreviations of the various column headings in Tables A.1 and A.4. The variable labels or abbreviations are listed in alphabetical order with detailed descriptions (where necessary) of how they are organized or obtained. Entries of (-1) under any of the table headings indicate missing data. Entries of (-2) indicate data from source catalogs that are incompatible

with the format presented in the tables.

A ACCELERATION AT SITE. Units of fraction of gravitational acceleration g.

ACCO ACCELERATION DATA CODE. Four Digit Binary Code, indicating basis of acceleration estimate in Table A.2.

0000 = Based on Joyner and Boore (1981) or Kawashima et al. (1984) attenuation relationships, as described in Section 3.4.2.

0001 = Based on attenuation relationship of site intensity data (e.g. Modified Mercali Scale), as provided by the source catalog or the source reference.

0010 = Based on response analysis, using a selected strong motion record, usually from another earthquake.

0100 = Based on strong motion record "nearby". Usually "nearby" implies several kilometers; or very close proximity, but is a bedrock record.

0110 = Based on "nearby" strong motion record and response analysis.

1000 = Based on strong motion record on site or at "nearby" sites with similar soil profile. Implies a great deal of precision associated with acceleration estimate.

ART ARTESIAN GROUND WATER CODE.

0 = not artesian; or artesian condition not noted.

1 = artesian condition noted.

BCAT BASE CATALOG CODE. Identifies which source catalog originally tabulated the data, or has formed the basis of other catalogs,

BCAT assists in reassembling any of the 7 original catalogs, if desired. See CAT for listing.

BORC GLOBAL BORING CODE. Identifies case studies using the exact same boring data. If BORC = 0, the boring is unique.

CASE CASE CODE. Consists of two parts. The first two digits identify the earthquake. The last two digits identify the case associated with that earthquake.

CAT SOURCE CATALOG CODE. Indicates data source catalog.

0 = Data not reported in previous catalogs

1 = Whitman (1971)

2 = Seed, Idriss and Arango (1975)

3 = Yegian (1976)

4 = Yegian and Vitelli (1981)

5 = Xie (1979)

6 = Davis and Berrill (1981)

7 = Tokimatsu and Yoshimi (1983)

8 = Seed, Tokimatsu, Harder and Chung (1984)

The following is used for ECAT only:

9 = Case study evlauation by writer (S. Liao)

CC CLAY CONTENT. Units of percent. For fines content FC below 10%, clay content CC assumed to be zero.

CSR CYCLIC SHEAR STRESS RATIO. In Table A.2, this quantity is calculated using the Seed and Idriss (1971) procedure. However, in Table A.4, CSR is calculated using slightly different procedures for those cases reported by Whitman (1971) and Tokimatsu and Yoshimi (1983).

CSRN CYCLIC SHEAR STRESS RATIO, NORMALIZED.

CSR values normalized for earthquake magnitude according to Seed et al. (1984). See section 3.4.2 for details.

DC DISTANCE CODE. Two digit binary code indicating classifications of distance measures EP and DER

1st digit: 0 = EP based on location of maximum intensity

1 = EP is instrumental epicenter.

2nd digit: 0 = DER assumed the same as EP

1 = DER estimated independently of EP.

DER DISTANCE TO ENERGY RELEASE. Units of kilometers. Closest distance to fault rupture or to the zone of major energy release. A distance measure that is largely judgmental in many cases. If a measure of DER is not available, DER is set to EP.

DOC DOCUMENTATION CODE. Classifies the amount of documentation on soil data available.

1 = Well documented. Boring logs available, with detailed soil descriptions.

2 = Adequately documented. Boring logs not available. Soil data presented usually in the form of a generalized profile.

3 = Poorly documented. Usually only the SPT N-value is given. Boring logs and soil profile descriptions not given.

D50 MEDIAN GRAIN SIZE. Units of millimeters.

DUR DURATION OF SHAKING. Units of seconds.

ECAT — ENHANCED CATALOG CODE. Identified which source catalog provided most of data (primarily the soil data) reported in Table A.2. See CAT for listing.

EP — EPICENTRAL DISTANCE TO SITE. Units of kilometers. In many cases, only DER is given and EP is assumed equal to DER in Table A.2.

FC — FINES CONTENT. Units of percent. Fines defined as material passing No. 200 sieve.

GC — GRAVEL CONTENT. Units of percent. Gravel defined as material retained on No. 4 sieve.

GRAD — GRADATION CODE. Four digit binary code based on classifications of gradation measures FC, GC, D50, UC, or based on non quantitative descriptions (e.g., clean sand or silty sand).

1st digit: 0 = $D_{50} > 0.25$

1 = $D_{50} < 0.25$, or soil described as very fine sand.

2nd digit: 0 = PF ≤ 12

1 = PF > 12, or soil described as silty or clayey.

3rd digit: 0 = GC ≤ 10%

1 = GC > 10%, or soil described as gravelly.

4th digit: 0 = UC ≤ 6, or soil described as poorly graded or predominantly one or two size categories, e.g. fine to medium sand.

1 = UC > 6, or soil described as well-graded, or having a wide range of grain sizes.

H FOCAL DEPTH. Units of kilometers.

HY HYPOCENTRAL DISTANCE TO SITE. Units of kilometers.

IE ith MOST RECENT EARTHQUAKE AT THE SITE. (See NE)

IL ith MOST RECENT TIME SITE HAS LIQUEFIED. (See NL)

IN ith MOST RECENT TIME SITE HAS NOT LIQUEFIED. (See NN)

LIQ1 LIQUEFACTION CODE 1.

0. = no liquefaction

1. = liquefaction

LIQ2 LIQUEFACTION CODE 2. Based on Tokimatsu & Yoshimi (1983).

.0 = no liquefaction

.5 = marginal site, defined as a site located just outside the boundary of a liquefaction zone

.7 = moderate liquefaction, indicated by appearance of sand boils, but with minor ground or foundation movements.

1.0 = Extensive liquefaction, indicated by sand boils and/or major ground or foundation movements.

M RICHTER MAGNITUDE.

N SPT RESISTANCE.

N1 SPT RESISTANCE MODIFIED. In Table A.3, the value reported for N1 is the value reported in the source catalog, except for Catalog 3 -Yegian (1976). For N1 values reported by catalog 3, the correction factor used is that proposed by Liao and Whitman (1986) -See Sec. 3.4.3. All of the values of N1 in Table A.2 are calculated based on the Liao and Whitman correction factor.

NAV N-VALUE AVERAGE CODE. Three digit binary code indicating method of selecting a representative N-value.

000 = Method unknown or abitrary

100 = Critical N value, either minimum N_1 value or obtained using Seed and Idriss (1971) scheme

010 = Layer average, usually involving one boring.

001 = Site average, usually obtained in a layer, but involving more than one boring.

NC1, NC2 NUMBER OF CORRELATED BORINGS. Estimates of the number of borings that could be considered as being at the same site. Usually the case studies for these borings are identified with the same site name. NC1 is a lower bound estimate and NC2 is an upper bound estimate.

NE NUMBER OF EARTHQUAKE SHAKINGS. Number of times a given site (boring location) is used in different case studies for a series of earthquakes in the same geographic vicinity. (See IE)

NL NUMBER OF LIQUEFACTION OCCURRENCES. Number of times a site (boring location) has been reordered as experiencing liquefaction. (See IL)

NLG NOT-LEVEL GROUND CODE.

0 = level ground although possibly with buildings and foundations nearby.

1 = not level ground: slope, dike, dam or embankment.

NN NUMBER OF NON-LIQUEFACTION OCCURRENCES. Number of times a site (boring location) has been recorded as not liquefying during various earthquakes. (See IN)

REG — REGION CODE. Two digit binary code indicating geographic region

100 = Japan

010 = California

001 = China

000 = Other Region

REP — REPORTING CATALOGS CODE. Identifies which catalogs have previously included a particular case. Seven digit binary code. Each digit is either a "0" indicating "no", or a "1" indicating "yes". The position of each digit corresponds to a certain published catalog, as referenced by CAT:

Digit Position	Catalogue
1	Whitman (1971)
2	Seed, Idriss and Arango (1975)
3	Yegian (1976)
4	Yegian and Vitelli (1981)
5	Xie (1979)
6	Davis and Berrill (1981)
7	Tokimatsu and Yoshimi (1983)
8	Seed, Tokimatsu, Harder, and Chung (1984)

EXAMPLE: REP = 1000100 indicates the case was reported in catalogs by Whitman (1971) and Xie (1979).

SIGE — EFFECTIVE VERTICAL STRESS. Units of kg/cm^2.

SIGT — TOTAL VERTICAL STRESS. Units of kg/cm^2.

SITE GLOBAL SITE CODE. Identifies sites in close geographical proximity to each other, usually identified with the same name, usually expected to experience the same ground motions during a given earthquake. If SITE = 0, this site is unique. Some judgment in assigning this code based on whether sites could be considered independent cases or not.

SPTC SPT METHOD CODE. Four digit binary code indicating method of obtaining SPT data.

0000 = SPT, method unknown

1000 = SPT, conventional rope and drum

0100 = SPT, free fall hammer

0010 = N-value estimated from static cone

0001 = N-value obtained by method other than SPT or cone, e.g. non-standard dyanmic penetration test or based on relative density measurements.

TAB TABLUATED CATALOG CODE. Identifies which catalog's data are reported in Table 4, for a particular case. Eight Digit Binary Code. Each digit is either a "1" , indicating "yes" or a "0", indicating the portion of each digit corresponds to a certain published source catalog, as referenced by CAT. Same as digit listing for REP.

TB TIME OF BORING CODE. Two digit binary code, indicating when boring was performed.

10 = Before earthquake

01 = After earthquake

00 = Unknown time

ZL CRITICAL DEPTH OF LIQUEFACTION. Units of meters.

ZW DEPTH TO WATER TABLE. Units of meters.

Table A.1

Case Identification and Qualitative Variables

ND	EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	REG	DC	ACCO	SPTC	NAV	NLG	ART	TB	NC1	NC2	NE	IE	NL	IL	NN	IN	DOC
2																								
	Niigata	1802	Niigata	0101	0	1	1	100	00	0001	0000	001	0	0	01	1	2	3	3	1	0	2	2	2
	Sado Island	12/1	Niigata	0102	0	1	2	100	00	0001	0000	001	0	0	01	1	2	3	3	0	0	3	3	2
2																								
	Niigata	1877	Niigata	0201	0	1	1	100	00	0001	0000	001	0	0	01	1	2	3	2	1	0	2	1	2
	Koshigun	7/22	Niigata	0202	0	1	2	100	00	0001	0000	001	0	0	01	1	2	3	2	0	0	3	2	2
6																								
	Mino-Owari	1891	Ogaki	0301	1	0	0	100	00	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1
		10/28	Ginan	0302	1	14	14	100	00	0001	0100	100	0	0	01	1	1	2	1	2	2	0	0	1
			Unuma	0303	1	0	0	100	00	0001	0100	100	0	1	01	1	1	1	1	1	1	0	0	1
			Ogase Pond	0304	1	0	0	100	00	0001	0100	100	0	1	01	1	1	1	1	1	1	0	0	1
			Saya	0306	1	4	4	100	00	0000	0000	100	0	0	01	1	1	2	2	1	1	0	0	1
			Biwajima	0307	1	5	5	100	00	0000	0000	100	0	0	01	1	1	2	2	1	1	0	0	1
5																								
	Tokyo	1894	Tone River	0401	1	8	8	100	00	0000	0000	100	0	0	01	1	1	3	3	2	2	1	0	1
		6/20	Gyoda	0402	0	9	9	100	00	0000	0000	100	0	0	01	1	1	3	3	1	0	2	2	1
			Kasu	0403	0	10	10	100	00	0000	0000	100	0	0	01	1	1	3	3	0	0	3	3	1
			Kasukabe	0404	1	11	11	100	00	0000	0000	100	0	0	01	1	1	3	3	3	3	0	0	1
			Ara River	0405	1	12	12	100	00	0000	0000	100	0	0	01	1	1	3	2	2	1	1	0	1
6																								
	San Francisco	1906	Foot of Market-b	0502	1	3	18	010	00	0010	0001	100	0	0	01	2	2	2	2	2	2	0	0	3
			So. of Market	0504	1	20	0	010	00	0010	1000	001	0	0	01	1	1	2	2	2	2	0	0	3
			Mission Creek	0505	1	21	0	010	00	0001	1000	001	0	0	01	1	1	2	2	2	2	0	0	1
			Salinas	0506	1	0	0	010	01	0000	1000	001	0	0	01	1	1	1	1	1	1	0	0	1
			Yerba Buena Cove	0507	1	3	3	010	00	0010	0100	001	0	0	01	2	2	3	3	1	1	2	0	2
			Telegraph Hill	0508	0	32	32	010	00	0010	0100	001	0	0	01	1	1	3	3	0	0	3	3	2
2																								
	Gono	1909	Saya	0601	0	4	4	100	00	0000	0000	100	0	0	01	1	1	2	1	1	0	1	1	1
		8/14	Biwajima	0602	0	5	5	100	00	0000	0000	100	0	0	01	1	1	2	1	1	0	1	1	1
12																								
	Kanto	1923	Arakawa 7	0702	1	6	6	100	00	0001	0000	000	0	0	01	1	6	1	1	1	1	0	0	3
		9/1	Arakawa 12	0703	1	6	0	100	00	0001	0000	000	0	0	01	1	6	1	1	1	1	0	0	3
			Arakawa 21	0704	1	6	0	100	00	0001	0000	000	0	0	01	1	6	1	1	1	1	0	0	3
			Arakawa 30	0705	1	6	0	100	00	0001	0000	000	0	0	01	1	6	1	1	1	1	0	0	3
			Arakawa 49	0707	0	6	7	100	00	0001	0000	000	0	0	01	1	6	1	1	0	0	1	1	3
			Tone River	0708	0	8	8	100	00	0000	0000	100	0	0	01	1	1	3	2	2	0	1	1	1
			Gyoda	0709	0	9	9	100	00	0000	0000	100	0	0	01	1	1	3	2	1	0	2	1	1
			Kasu	0710	0	10	10	100	00	0000	0000	100	0	0	01	1	1	3	2	1	0	2	2	1
			Kasukabe	0711	1	11	11	100	00	0000	0000	100	0	0	01	1	1	3	2	2	2	0	0	1
			Ara River	0712	1	12	12	100	00	0000	0000	100	0	0	01	1	6	3	2	1	1	0	0	1
			Ukita	0713	1	0	0	100	00	0000	0000	100	0	0	01	1	1	1	1	1	1	0	0	1
			Edogawa	0714	1	0	0	100	00	0000	0000	100	0	0	01	1	1	1	1	1	1	0	0	1
1																								
	Santa Barbara	1925	Sheffield Dam	0801	1	0	0	010	00	0010	0001	001	1	0	10	1	1	1	1	1	1	0	0	2
5																								
	Nishi-Saitama	1931	Tone River	0901	1	8	8	100	00	0000	0000	100	0	0	01	1	1	3	1	2	1	1	0	1
		9/21	Gyoda	0902	1	9	9	100	00	0000	0000	100	0	0	01	1	1	3	1	1	1	2	0	1
			Kasu	0903	0	10	10	100	00	0000	0000	100	0	0	01	1	1	3	1	0	0	3	1	1
			Kasukabe	0904	1	11	11	100	00	0000	0000	100	0	0	01	1	1	3	1	3	1	0	0	1
			Ara River	0905	0	12	12	100	00	0000	0000	100	0	0	01	1	1	3	1	2	0	1	1	1
5																								
	Long Beach	1933	LNG Ter./Res. Pt-1	1001	0	13	0	010	01	0110	1000	100	0	0	01	4	4	1	1	0	0	1	1	1
		3/10	LNG Ter./Res. Pt-2	1002	0	13	0	010	01	0110	1000	100	0	0	01	4	4	1	1	0	0	1	1	1
			LNG Ter./Res. Pt-3	1003	0	13	0	010	01	0110	1000	100	0	0	01	4	4	1	1	0	0	1	1	1
			LNG Ter./Res. Pt-4	1004	0	13	0	010	01	0110	1000	100	0	0	01	4	4	1	1	0	0	1	1	1
			L.A. Pier A	1005	0	0	0	010	01	0110	1000	001	0	0	01	1	1	1	1	0	0	1	1	2
3																								
	El Centro	1940	Brawley	1101	1	0	0	100	00	0001	0001	001	0	0	10	1	1	1	1	1	1	0	0	1
			All-Am. Canal	1102	1	0	0	100	00	0001	0001	001	0	0	10	1	1	1	1	1	1	0	0	1
			Solfatara Canal	1103	1	0	0	100	00	0001	0001	001	0	0	10	1	1	1	1	1	1	0	0	1
4																								
	Tonankai	1944	Komei	1201	1	0	0	100	01	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1
		12/7	Meiko St.	1202	1	0	0	100	01	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1

Table A.1 (Continued)

ND	EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	REG	DC	ACCO	SPTC	NAV	NLG	ART	TB	NC1	NC2	NE	IE	NL	IL	NN	IN	DOC
			Ienaga	1203	1	0	0	100	01	0001	0100	010	0	0	01	1	1	1	1	1	1	0	0	1
			Ginan	1204	1	14	14	100	01	0001	0100	010	0	0	01	1	1	2	1	2	1	0	0	1
1	Fukui	1948	Shonenji	1304	1	0	0	100	00	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1
1	San Francisco	1955	Joaquin Aqueduct	1401	0	29	0	010	00	0000	0001	100	0	0	01	1	1	2	2	0	0	2	2	3
32	San Francisco	1957	St. Francis Circle	1501	0	0	0	010	00	0000	0001	100	0	0	00	1	1	1	1	0	0	1	1	3
	Daly City	3/22	Lake Merced	1502	1	0	0	010	00	0000	1000	001	1	0	00	1	1	1	1	1	1	0	0	3
			Duboce & Sanchez	1503	0	0	0	010	00	0000	0001	100	0	0	00	1	1	1	1	0	0	1	1	3
			Foot of Market-b	1505	0	3	18	010	00	0000	0001	100	0	0	00	2	2	2	1	1	0	1	1	3
			So. of Market	1507	0	20	0	010	00	0000	0001	100	0	0	00	1	1	2	1	1	0	1	1	3
			Mission Creek	1508	0	21	0	010	00	0000	0001	100	0	0	00	1	1	2	1	1	0	1	1	3
			Polk & Golden Gate	1509	0	0	0	010	00	0000	0001	100	0	0	00	1	1	1	1	0	0	1	1	3
			Polk & Market	1510	0	0	0	010	00	0000	0001	100	0	0	00	1	1	1	1	0	0	1	1	3
			Welden-a	1511	0	22	0	010	00	0000	1000	100	0	0	00	3	3	1	1	0	0	1	1	3
			Welden-b	151[illegible]	0	22	0	010	00	0000	1000	100	0	0	00	3	3	1	1	0	0	1	1	3
			Welden-d	1514	0	22	0	010	00	0000	1000	100	0	0	00	3	3	1	1	0	0	1	1	3
			Mission & Spear-a	1515	0	23	0	010	00	0000	1000	100	0	0	00	2	2	1	1	0	0	1	1	3
			Mission & Spear-b	1516	0	23	0	010	00	0000	1000	100	0	0	00	2	2	1	1	0	0	1	1	3
			Park & Otis,Al.-a	1517	0	24	0	010	00	0000	1000	100	0	0	00	2	2	1	1	0	0	1	1	3
			Park & Otis,Al.-b	1518	0	24	0	010	00	0000	1000	100	0	0	00	2	2	1	1	0	0	1	1	3
			Singleton,Alemeda	1519	0	0	0	010	00	0000	1000	100	0	0	00	1	1	1	1	0	0	1	1	3
			Treasure Island-a	1520	0	25	0	010	00	0000	1000	100	0	0	00	8	8	1	1	0	0	1	1	3
			Treasure Island-b	1521	0	25	0	010	00	0000	1000	100	0	0	00	8	8	1	1	0	0	1	1	3
			Treasure Island-c	1522	0	25	0	010	00	0000	1000	100	0	0	00	8	8	1	1	0	0	1	1	3
			Treasure Island-d	1523	0	25	0	010	00	0000	1000	100	0	0	00	8	8	1	1	0	0	1	1	3
			Treasure Island-f	1525	0	25	0	010	00	0000	1000	100	0	0	00	8	8	1	1	0	0	1	1	3
			Treasure Island-g	1526	0	25	0	010	00	0000	1000	100	0	0	00	8	8	1	1	0	0	1	1	3
			Treasure Island-h	1527	0	25	0	010	00	0000	1000	100	0	0	00	8	8	1	1	0	0	1	1	3
			Treasure Island-i	1528	0	25	0	010	00	0000	1000	100	0	0	00	8	8	1	1	0	0	1	1	3
			W 5th/Ave D,Al.-a	1529	0	26	0	010	00	0000	1000	100	0	0	00	2	2	1	1	0	0	1	1	3
			W 5th/Ave D,Al.-b	1530	0	26	0	010	00	0000	1000	100	0	0	00	2	2	1	1	0	0	1	1	3
			Westline Ave,Al.	1531	0	0	0	010	00	0000	1000	100	0	0	00	1	1	1	1	0	0	1	1	3
			Emeryville,Al.	1532	0	0	0	010	00	0000	1000	100	0	0	00	1	1	1	1	0	0	1	1	3
			Westline M.C.-a	1533	0	27	0	010	00	0000	0001	100	0	0	00	2	2	1	1	0	0	1	1	3
			Westline M.C.-b	1534	0	27	0	010	00	0000	0001	100	0	0	00	2	2	1	1	0	0	1	1	3
			Yerba Buena Cove	1535	0	3	3	010	00	0010	0100	001	0	0	01	2	2	3	2	1	0	2	2	2
			Telegraph Hill	1536	0	32	32	010	00	0010	0100	001	0	0	01	1	1	3	2	0	0	3	2	2
5	Chile	1960	Puerto Montt-a	1601	1	64	0	000	11	0001	1000	000	0	0	01	1	2	1	1	1	1	0	0	3
		5/22	Puerto Montt-b	1602	1	64	0	000	11	0001	1000	000	0	0	01	1	2	1	1	1	1	0	0	3
			Puerto Montt-c	1603	0	0	0	000	11	0001	1000	000	0	0	01	1	1	1	1	0	0	1	1	3
			Concepcion	1604	0	0	0	000	11	0001	1000	001	0	0	01	1	1	1	1	0	0	1	1	3
			Huachipato	1605	0	0	0	000	11	0001	1000	001	0	0	01	1	1	1	1	0	0	1	1	3
5	Alaska	1964	Snow River B605A	1701	1	65	65	000	11	0001	1000	001	0	0	00	1	2	1	1	1	1	0	0	1
		3/27	Snow River B605	1702	1	65	65	000	11	0001	1000	001	0	0	00	1	2	1	1	1	1	0	0	1
			Quartz Creek	1703	0	0	0	000	11	0001	1000	100	0	0	00	1	1	1	1	0	0	1	1	1
			Scott Glacier	1704	1	0	0	000	11	0001	1000	001	0	0	00	1	1	1	1	1	1	0	0	1
			Valdez	1705	1	0	0	000	11	0001	1000	001	1	0	00	1	1	1	1	1	1	0	0	1
11	Niigata	1964	Niigata	1804	0	1	0	100	00	1000	0000	001	0	0	10	1	1	1	1	0	0	1	1	2
		6/16	Niigata	1814	1	1	1	100	00	1000	0000	001	0	0	10	1	1	3	1	1	1	2	0	2
			Niigata	1815	0	1	2	100	00	1000	0000	001	0	0	10	1	1	3	1	0	0	2	1	2
			Niigata	1816	0	1	0	100	00	1000	0000	000	0	0	10	1	1	1	1	0	0	1	1	1
			Niigata	1817	1	1	0	100	00	1000	0000	000	0	0	10	1	1	1	1	1	1	0	0	1
			Niigata	1818	0	1	0	100	00	1000	0000	000	0	0	10	1	1	1	1	0	0	1	1	1
			Niigata	1819	0	1	0	100	00	1000	0000	000	0	0	10	1	1	1	1	0	0	1	1	1
			Showa Bridge 2	1820	1	0	0	100	00	1000	0000	000	0	0	10	1	1	1	1	1	1	0	0	3
			Showa Bridge 4	1821	0	0	0	100	00	1000	0000	000	0	0	01	1	1	1	1	0	0	1	1	3
			Road Site	1822	0	0	0	100	00	1000	0000	000	0	0	01	1	1	1	1	0	0	1	1	1
			River Site	1823	1	0	0	100	00	1000	0000	000	0	0	01	1	1	1	1	1	1	0	0	1
1	San Francisco	1965	Joaquin Aqueduct	1901	0	29	29	010	00	0000	0001	100	0	0	01	1	1	2	1	0	0	2	1	3

Table A.1 (Continued)

ND	EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	REG	DC	ACCD	SPTC	NAV	NLG	ART	TB	NC1	NC2	NE	IE	NL	IL	NN	IN	DOC
1																								
	Caracus	1967	Caraballeda	2001	1	0	0	000	00	0001	1000	000	0	0	01	1	1	1	1	1	1	0	0	3
7																								
	Tokachi-Oki	1968	Nanohama,Hakadote	2201	1	0	0	100	11	0001	0000	000	0	0	10	1	1	1	1	1	1	0	0	1
		5/16	Hachinohe P-1	2206	0	30	0	100	11	1000	0000	100	0	0	10	1	4	1	1	0	0	1	1	1
			Hachinohe P-2	2207	0	30	0	100	11	1000	0000	010	0	0	10	1	4	1	1	0	0	1	1	1
			Hachinohe P-4	2208	0	30	0	100	11	1000	0000	100	0	0	10	1	4	1	1	0	0	1	1	1
			Hachinohe P-5	2209	0	30	0	100	11	1000	0000	100	0	0	10	1	4	1	1	0	0	1	1	1
			Hachinohe P-6	2210	1	30	0	100	11	1000	0000	100	0	0	10	1	1	1	1	1	1	0	0	1
			Hachinohe Accel.	2212	0	0	0	100	11	1000	0000	100	0	0	10	1	1	1	1	0	0	1	1	1
8																								
	Saitama	1968	Saitama 101-2	2301	0	0	0	100	00	0000	0000	100	0	0	00	1	5	1	1	0	0	1	1	1
		7/1	Saitama 105-2	2302	0	0	0	100	00	0000	0000	100	0	0	00	1	5	1	1	0	0	1	1	1
			Saitama 119	2303	0	0	0	100	00	0000	0000	100	0	0	00	1	5	1	1	0	0	1	1	1
			Saitama 121	2304	0	0	0	100	00	0000	0000	100	0	0	00	1	5	1	1	0	0	1	1	1
			Saitama 130	2305	0	0	0	100	00	0000	0000	100	0	0	00	1	5	1	1	0	0	1	1	1
			Saitama 602	2306	0	0	0	100	00	0000	0000	100	0	0	00	1	1	1	1	0	0	1	1	1
2																								
	Santa Rosa	1969	Yerba Buena Cove	2401	0	3	3	010	10	0100	0100	001	0	0	01	1	1	3	1	1	0	2	1	2
		10/1	Telegraph Hill	2402	0	32	32	010	10	0100	0100	001	0	0	01	1	1	3	1	0	0	3	1	2
1																								
	Gediz,Turkey	1970	Bursa	2501	0	0	0	000	00	0010	1000	001	0	0	10	1	1	1	1	0	0	1	1	3
3																								
	San Fernando	1971	Juvenile Hall	2601	1	0	0	010	01	0100	1000	000	0	0	00	1	1	1	1	1	1	0	0	3
		2/9	Jensen Plant	2602	1	0	0	010	01	0100	1000	000	1	0	00	1	1	1	1	1	1	0	0	1
			Van Norman Dam	2603	1	0	0	010	01	0110	1000	000	1	0	00	1	1	1	1	1	1	0	0	2
1																								
	Yokohama	1972	Yokohama	2701	0	0	0	100	10	1000	0000	001	0	0	00	1	1	1	1	0	0	1	1	3
11																								
	Haicheng	1975	Shuang. Eardo Br.	2803	0	35	0	001	10	0001	0100	000	0	0	00	1	1	1	1	0	0	1	1	3
		2/4	Shenglitang	2805	0	0	0	001	10	0001	0100	000	0	0	00	1	1	1	1	0	0	1	1	3
			Ligohe Plant	2806	1	38	0	001	10	0001	0100	000	0	0	00	1	1	1	1	1	1	0	0	3
			Panjin Storage	2808	1	0	0	001	10	0001	0100	100	0	0	00	1	1	1	1	1	1	0	0	3
			Yinkou Paper Plant	2809	1	0	0	001	10	0001	0100	000	0	0	00	1	1	1	1	1	1	0	0	1
			Nanheyan Irr. Sta.	2810	1	0	0	001	10	0001	1000	000	0	0	00	1	1	1	1	1	1	0	0	3
			Shuiyuan Commune	2811	1	0	0	001	10	0001	0100	000	0	0	00	1	1	1	1	1	1	0	0	3
			Yinkou Gate	2812	1	0	0	001	10	0001	0100	000	0	0	00	1	1	1	1	1	1	0	0	3
			Panjin Ch. Fertil.	2813	1	0	0	001	10	0001	1000	100	0	0	00	1	1	1	1	1	1	0	0	1
			Yinkou Glass Fiber	2814	1	0	0	001	10	0001	0100	100	0	0	00	1	1	1	1	1	1	0	0	1
			Shuang Tai Zi R.	2815	0	34	0	001	10	0001	0100	100	0	0	00	1	1	1	1	0	0	1	1	1
4																								
	Guatemala	1976	Amatitlan 1	2901	1	0	0	000	01	0001	1000	100	0	0	00	1	1	1	1	1	1	0	0	1
		2/4	Amatitlan 2	2902	0	0	38	000	01	0001	1000	100	0	0	00	1	1	1	1	0	0	1	1	1
			Amatitlan 3	2903	0	37	0	000	01	0001	1000	100	0	0	00	2	2	1	1	0	0	1	1	1
			Amatitlan 4	2904	0	37	0	000	01	0001	1000	100	0	0	00	2	2	1	1	0	0	1	1	1
9																								
	Tangshan	1976	Weigezhuang	3001	1	0	0	001	10	0001	0100	000	0	0	00	1	1	1	1	1	1	0	0	3
		7/28	Lujiato Mine	3002	1	0	0	001	10	0001	0100	000	0	0	00	1	1	1	1	1	1	0	0	3
			Tangshan City	3006	0	0	0	001	10	0001	0100	100	0	0	01	1	1	1	1	0	0	1	1	1
			Qing Yin	3007	1	0	0	001	10	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1
			Le Ting	3009	1	0	0	001	10	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1
			Coastal Region	3010	1	0	0	001	10	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1
			Yao Yuan Village	3011	1	0	0	001	10	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1
			Ma Feng	3012	0	0	0	001	10	0001	0100	100	0	0	01	1	1	1	1	0	0	1	1	1
			Wang Zhuang	3013	1	0	0	001	10	0001	0100	100	0	0	01	1	1	1	1	1	1	0	0	1
17																								
	San Juan	1977	Barrio Castro B-1	3201	1	39	0	000	10	0100	1000	100	0	0	01	2	2	1	1	1	1	0	0	1
	Argentina	11/3	Barrio Castro B-2	3202	1	39	0	000	10	0100	1000	100	0	0	01	2	2	1	1	1	1	0	0	1
			Caucete B-3	3203	1	0	0	000	10	0100	1000	100	0	0	01	1	1	1	1	1	1	0	0	1
			West of River B-4	3204	0	0	0	000	10	0100	1000	100	0	0	01	1	1	1	1	0	0	1	1	1
			West of River B-5	3205	0	0	0	000	10	0100	1000	100	0	0	01	1	1	1	1	0	0	1	1	1
			Fin. Santiago B-6	3206	1	0	0	000	10	0100	1000	100	0	0	01	1	1	1	1	1	1	0	0	1
			Escuela Normal 1	3207	1	55	0	000	10	0100	1000	100	0	0	01	6	6	1	1	1	1	0	0	1
			Escuela Normal 2	3208	1	55	0	000	10	0100	1000	100	0	0	01	6	6	1	1	1	1	0	0	1
			Escuela Normal 3	3209	1	55	0	000	10	0100	1000	100	0	0	01	6	6	1	1	1	1	0	0	1
			Escuela Normal 4	3210	1	55	0	000	10	0100	1000	100	0	0	01	6	6	1	1	1	1	0	0	1
			Escuela Normal 5	3211	1	55	0	000	10	0100	1000	100	0	0	01	6	6	1	1	1	1	0	0	1

Table A.1 (Continued)

ND	EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	REG	DC	ACCO	SPTC	NAV	NLG	ART	TB	NC1	NC2	NE	IE	NL	IL	NN	IN	DOC
			Escuela Normal 6	3212	1	55	0	000	10	0100	1000	100	0	0	01	6	6	1	1	1	1	0	0	1
			Airport (Rt 20) 1	3213	0	56	0	000	10	0100	1000	100	0	0	01	2	2	1	1	0	0	1	1	1
			Airport (Rt 20) 2	3214	0	56	0	000	10	0100	1000	100	0	0	01	2	2	1	1	0	0	1	1	1
			Santa Rosa 1	3215	0	57	0	000	10	0100	1000	100	0	0	01	3	3	1	1	0	0	1	1	1
			Santa Rosa 2	3216	0	57	0	000	10	0100	1000	100	0	0	01	3	3	1	1	0	0	1	1	1
			Santa Rosa 3	3217	0	57	0	000	10	0100	1000	100	0	0	01	3	3	1	1	0	0	1	1	1
1	Izu	1978	Mochikoshi	3301	1	0	0	100	10	0100	0010	001	1	0	01	1	1	1	1	1	1	0	0	2
25	Miyagiken-Oki - 1	1978 2/20	Arahama	3401	0	49	0	100	10	0100	0000	000	0	0	01	1	1	2	2	1	0	1	1	3
			Nakamura 1	3402	0	40	0	100	10	0100	0000	000	0	0	01	1	1	2	2	0	0	2	2	3
			Nakamura 4	3403	1	40	0	100	10	0100	0000	000	0	0	01	1	2	2	2	2	1	0	1	3
			Nakamura 5	3404	0	40	0	100	10	0100	0000	000	0	0	01	1	2	2	2	1	0	1	1	3
			Yuriageka 1	3405	0	41	0	100	10	0100	0000	000	0	0	01	1	2	2	2	1	0	1	1	3
			Yuriageka 2	3406	0	41	0	100	10	0100	0000	000	0	0	01	1	2	2	2	1	0	1	1	3
			Yuriageka 3	3407	0	41	0	100	10	0100	0000	000	0	0	01	1	1	2	2	1	0	2	2	3
			Yuriage Bridge 1	3408	0	42	0	100	10	0100	0000	000	0	0	01	1	3	2	2	1	0	1	1	3
			Yuriage Bridge 2	3409	0	42	0	100	10	0100	0000	000	0	0	01	1	3	2	2	1	0	1	1	3
			Yuriage Bridge 3	3410	0	42	0	100	10	0100	0000	000	0	0	01	1	3	2	2	1	0	1	1	3
			Yuriage Bridge 5	3411	0	42	0	100	10	0100	0000	000	0	0	01	1	1	2	2	0	0	2	2	3
			Oiira 1	3412	0	43	0	100	10	0100	0000	000	0	0	01	1	2	2	2	0	0	2	2	3
			Oiira 2	3413	0	43	0	100	10	0100	0000	000	0	0	01	1	2	2	2	1	0	1	1	3
			Kitawabu 2	3414	0	44	0	100	10	0100	0000	000	1	0	01	1	2	2	2	1	0	1	1	3
			Kitawabu 3	3415	0	44	0	100	10	0100	0000	000	1	0	01	1	1	2	2	0	0	2	2	3
			Shiomi 2	3416	0	45	0	100	10	0100	0000	000	0	0	10	1	1	2	2	0	0	2	2	3
			Shiomi 6	3417	0	45	0	100	10	0100	0000	000	0	0	10	1	1	2	2	1	0	1	1	3
			Hiyori 5	3418	0	46	0	100	10	0100	0000	000	0	0	01	1	1	2	2	0	0	2	2	3
			Hiyori 18	3419	0	46	0	100	10	0100	0000	000	0	0	01	1	1	2	2	1	0	1	1	3
			Nakajima 2	3420	0	47	0	100	10	0100	0000	000	0	0	01	1	1	2	2	0	0	2	2	3
			Nakajima 18	3421	0	47	0	100	10	0100	0000	000	0	0	01	1	1	2	2	1	0	1	1	3
			Sendaikou 1	3422	0	48	0	100	10	1000	0000	000	0	0	01	1	2	2	2	0	0	2	2	3
			Sendaikou 4	3423	0	48	0	100	10	1000	0000	000	0	0	01	1	2	2	2	0	0	2	2	3
			Ishinomaki 2	3424	0	50	50	100	10	0100	0000	010	0	0	10	1	1	2	2	1	0	1	1	3
			Ishinomaki 4	3425	0	50	51	100	10	0100	0000	100	0	0	10	1	1	2	2	0	0	2	2	3
25	Miyagiken-Oki - 2	1978 6/12	Arahama	3501	1	49	0	100	11	0100	0000	000	0	0	01	1	1	2	1	1	1	1	0	3
			Nakamura 1	3502	0	40	0	100	11	0100	0000	000	0	0	01	1	1	2	1	0	0	2	1	3
			Nakamura 4	3503	1	40	0	100	11	0100	0000	000	0	0	01	1	2	2	1	2	1	1	0	3
			Nakamura 5	3504	1	40	0	100	11	0100	0000	000	0	0	01	1	2	2	1	1	1	1	0	3
			Yuriageka 1	3505	1	41	0	100	11	0100	0000	000	0	0	01	1	2	2	1	1	1	1	0	3
			Yuriageka 2	3506	1	41	0	100	11	0100	0000	000	0	0	01	1	2	2	1	1	1	1	0	3
			Yuriageka 3	3507	0	41	0	100	11	0100	0000	000	0	0	01	1	1	2	1	1	0	2	1	3
			Yuriage Bridge 1	3508	1	42	0	100	11	0100	0000	000	0	0	01	1	3	2	1	1	1	1	0	3
			Yuriage Bridge 2	3509	1	42	0	100	11	0100	0000	000	0	0	01	1	3	2	1	1	1	1	0	3
			Yuriage Bridge 3	3510	1	42	0	100	11	0100	0000	000	0	0	01	1	3	2	1	1	1	1	0	3
			Yuriage Bridge 5	3511	0	42	0	100	11	0100	0000	000	0	0	01	1	1	2	1	0	0	2	1	3
			Oiira 1	3512	1	43	0	100	11	0100	0000	000	0	0	01	1	2	2	1	2	1	1	0	3
			Oiira 2	3513	1	43	0	100	11	0100	0000	000	0	0	01	1	2	2	1	1	1	1	0	3
			Kitawabu 2	3514	1	44	0	100	11	0100	0000	000	1	0	01	1	1	2	1	1	1	1	0	3
			Kitawabu 3	3515	0	44	0	100	11	0100	0000	000	1	0	01	1	1	2	2	0	0	2	1	3
			Shiomi 2	3516	0	45	0	100	11	0100	0000	000	0	0	10	1	1	2	1	0	0	2	1	3
			Shiomi 6	3517	1	45	0	100	11	0100	0000	000	0	0	10	1	1	2	1	1	1	1	0	3
			Hiyori 5	3518	0	46	0	100	11	0100	0000	000	0	0	01	1	1	2	1	0	0	2	1	3
			Hiyori 18	3519	1	46	0	100	11	0100	0000	000	0	0	01	1	1	2	1	1	1	1	0	3
			Nakajima 2	3520	0	47	0	100	11	0100	0000	000	0	0	01	1	1	2	1	0	0	2	1	3
			Nakajima 18	3521	1	47	0	100	11	0100	0000	000	0	0	01	1	1	2	1	1	1	1	0	3
			Sendaikou 1	3522	0	48	0	100	11	1000	0000	000	0	0	01	1	2	2	1	0	0	2	1	3
			Sendaikou 4	3523	0	48	0	100	11	1000	0000	000	0	0	01	1	2	2	1	0	0	2	1	3
			Ishinomaki 2	3524	1	50	50	100	11	0100	0000	010	0	0	10	1	1	2	1	1	1	1	0	3
			Ishinomaki 4	3525	0	50	51	100	11	0100	0000	100	0	0	10	1	1	2	1	0	0	2	1	3
2	Thessaloniki	1978 6/20	Greek Church	3601	0	0	0	000	10	0100	1000	001	0	0	10	1	1	1	1	0	0	1	1	2
			White Tower	3602	0	0	0	000	10	0100	1000	001	0	0	10	1	1	1	1	0	0	1	1	2
2	Guerrero	1979 3/14	Enmedio Zone 1	3701	0	0	0	000	10	0100	1000	001	0	0	10	1	1	1	1	0	0	1	1	2
			Enmedio Zone 2	3702	1	0	0	000	10	0100	1000	001	0	0	01	1	1	1	1	1	1	0	0	2

Table A.1 (Continued)

ND	EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	REG	DC	ACCO	SPTC	NAV	NLG	ART	TB	NC1	NC2	NE	IE	NL	IL	NN	IN	DOC
1																								
24	Montenegro	1979	Boca Kotorska	3801	1	0	0	000	10	0100	1000	010	0	0	10	1	1	1	1	1	1	0	0	1
	Imperial Valley	1979 10/15	Heber Rd. 1/A1	3901	0	52	79	010	11	0100	0100	001	0	0	01	1	1	2	2	0	0	2	2	1
			Heber Rd. 4.6/A2	3902	1	52	80	010	11	0100	0100	001	0	0	01	1	1	2	2	1	1	1	0	1
			Heber Rd. 7/A3	3903	0	52	81	010	11	0100	0100	001	0	0	01	1	1	2	2	0	0	2	2	1
			River Park A	3904	1	53	82	010	11	0100	1000	001	0	0	01	1	1	2	2	1	1	1	0	1
			River Park C	3905	1	53	83	010	11	0100	1000	001	0	0	01	1	1	2	2	1	1	1	0	1
			Wildlife 1Ns	3906	0	58	58	010	11	0100	0100	100	0	0	01	6	6	2	2	1	0	1	1	1
			Wildlife 2Ng1	3907	0	58	59	010	11	0100	0100	100	0	0	01	6	6	2	2	1	0	1	1	1
			Wildlife 2Ng3	3908	0	58	60	010	11	0100	0100	100	0	0	01	6	6	2	2	1	0	1	1	1
			Wildlife 3Ns	3909	0	58	61	010	11	0100	0100	100	0	0	01	6	6	2	2	1	0	1	1	1
			Wildlife 3Ns	3910	0	58	62	010	11	0100	0100	100	0	0	01	6	6	2	2	1	0	1	1	1
			Wildlife 5Ng	3911	0	58	63	010	11	0100	0100	100	0	0	01	6	6	2	2	1	0	1	1	1
			Vail V2a-b	3912	0	59	64	010	11	0100	0100	100	0	0	01	2	2	2	2	1	0	1	1	1
			Vail TV2a-b	3913	0	59	65	010	11	0100	0100	100	0	0	01	2	2	2	2	0	0	2	2	1
			Kornbloom K3	3914	0	60	66	010	11	0100	0100	100	0	0	01	4	4	2	2	1	0	1	1	1
			Kornbloom TK4a	3915	0	60	67	010	11	0100	0100	100	0	0	01	4	4	2	2	1	0	1	1	1
			Kornbloom TK4b	3916	0	60	68	010	11	0100	0100	100	0	0	01	4	4	2	2	1	0	1	1	1
			Kornbloom SK4-5	3917	0	60	69	010	11	0100	0100	100	0	0	01	4	4	2	2	1	0	1	1	1
			Radio Tower R2	3918	1	61	70	010	11	0100	0100	100	0	0	01	2	2	2	2	2	2	0	0	1
			Radio Tower R3	3919	0	61	71	010	11	0100	0100	100	0	0	01	2	2	2	2	0	0	2	2	1
			McKim TM6-7	3920	1	62	74	010	11	0100	0100	100	0	0	01	2	2	2	2	1	1	1	0	1
			McKim SM7	3921	1	62	75	010	11	0100	0100	100	0	0	01	2	2	2	2	1	1	1	0	1
			SNorthend SN2a	3922	0	63	76	010	11	0100	0100	100	0	0	01	2	2	2	2	0	0	2	2	1
			SNorthend SN2b	3923	0	63	77	010	11	0100	0100	100	0	0	01	2	2	2	2	0	0	2	2	1
			Young Y5	3924	0	0	78	010	11	0100	0100	100	0	0	01	1	1	2	2	0	0	2	2	1
1																								
2	Mexicali Val.	1980	Town of Delta	4001	1	0	0	010	11	0000	1000	100	0	0	01	1	1	1	1	1	1	0	0	2
	Mid-Chiba	1980	Owi Island 1	4101	0	54	54	100	10	1000	0000	100	0	0	10	1	1	1	1	0	0	1	1	1
		9/25	Owi Island 1	4103	0	54	54	100	10	1000	0000	100	0	0	10	1	1	1	1	0	0	1	1	1
24																								
	Westmoreland	1981 4/26	Heber Rd. 1/A1	4201	0	52	79	010	11	0100	0100	001	0	0	10	1	1	2	1	0	0	2	1	1
			Heber Rd. 4.6/A2	4202	0	52	80	010	11	0100	0100	001	0	0	10	1	1	2	1	1	0	1	1	1
			Heber Rd. 7/A3	4203	0	52	81	010	11	0100	0100	001	0	0	10	1	1	2	1	0	0	2	1	1
			River Park A	4204	0	53	82	010	11	0100	0100	001	0	0	10	1	1	2	1	1	0	1	1	1
			River Park C	4205	0	53	83	010	11	0100	0100	001	0	0	10	1	1	2	1	1	0	1	1	1
			Wildlife 1Ns	4206	1	58	58	010	11	0100	0100	100	0	0	01	6	6	2	1	1	1	1	1	1
			Wildlife 2Ng1	4207	1	58	59	010	11	0100	0100	100	0	0	01	6	6	2	1	1	1	1	1	1
			Wildlife 2Ng3	4208	1	58	60	010	11	0100	0100	100	0	0	01	6	6	2	1	1	1	1	1	1
			Wildlife 3Ns	4209	1	58	61	010	11	0100	0100	100	0	0	01	6	6	2	1	1	1	1	1	1
			Wildlife 3Ns	4210	1	58	62	010	11	0100	0100	100	0	0	01	6	6	2	1	1	1	1	1	1
			Wildlife 5Ng	4211	1	58	63	010	11	0100	0100	100	0	0	01	6	6	2	1	1	1	1	1	1
			Vail V2a-b	4212	0	59	64	010	11	0100	0100	100	0	0	01	2	2	2	1	0	0	2	1	1
			Vail TV2a-b	4213	0	59	65	010	11	0100	0100	100	0	0	01	2	2	2	1	0	0	2	1	1
			Kornbloom K3	4214	1	60	66	010	11	0100	0100	100	0	0	01	4	4	2	1	1	1	1	1	1
			Kornbloom TK4a	4215	1	60	67	010	11	0100	0100	100	0	0	01	4	4	2	1	1	1	1	1	1
			Kornbloom TK4b	4216	1	60	68	010	11	0100	0100	100	0	0	01	4	4	2	1	1	1	1	1	1
			Kornbloom SK4-5	4217	1	60	69	010	11	0100	0100	100	0	0	01	4	4	2	1	1	1	1	1	1
			Radio Tower R2	4218	1	61	70	010	11	0100	0100	100	0	0	01	2	2	2	1	2	1	0	0	1
			Radio Tower R3	4219	0	61	71	010	11	0100	0100	100	0	0	01	2	2	2	1	0	0	2	1	1
			McKim TM6-7	4220	0	62	74	010	11	0100	0100	100	0	0	01	2	2	2	1	1	0	1	1	1
			McKim SM7	4221	0	62	75	010	11	0100	0100	100	0	0	01	2	2	2	1	1	0	1	1	1
			SNorthend SN2a	4222	0	63	76	010	11	0100	0100	100	0	0	01	2	2	2	1	0	0	2	1	1
			SNorthend SN2b	4223	0	63	77	010	11	0100	0100	100	0	0	01	2	2	2	1	0	0	2	1	1
			Young Y5	4224	0	0	78	010	11	0100	0100	100	0	0	01	1	1	2	1	0	0	2	1	1

Table A.2

Enhanced/Edited Data

CASE	ECAT	LIQ1	LIQ2	M	H	EP	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC	GRAD
2																									
0101	8	0.	.0	6.6	-1.	39.	39.	20.	.12	.13	.12	1.0	7.0	1.35	.74	8.	9.3	1.09	10.1	2.	0.	0.	.3	2.5	0000
0102	8	0.	.0	6.6	-1.	39.	39.	20.	.12	.13	.12	1.0	7.0	1.35	.74	12.	13.9	1.09	15.2	2.	0.	0.	.3	2.5	0000
2																									
0201	8	0.	.0	6.1	-1.	47.	47.	12.	.08	.09	.07	1.0	7.0	1.35	.74	8.	9.3	1.09	10.1	2.	0.	0.	.3	2.5	0000
0202	8	0.	.0	6.1	-1.	47.	47.	12.	.08	.09	.07	1.0	7.0	1.35	.74	12.	13.9	1.09	15.2	2.	0.	0.	.3	2.5	0000
6																									
0301	2	1.	1.0	7.9	-1.	30.	30.	75.	.32	.33	.35	.8	13.7	2.59	1.30	17.	14.9	1.30	19.4	5.	0.	5.	.5	3.5	0000
0302	2	1.	1.0	7.9	-1.	30.	30.	75.	.32	.33	.35	2.0	9.1	1.73	1.02	10.	10.1	1.17	11.8	5.	0.	0.	.25	2.2	0000
0303	8	1.	1.0	7.9	-1.	30.	30.	75.	.32	.31	.33	1.9	6.0	1.17	.75	17.	19.8	1.30	25.5	3.	0.	25.	.8	3.4	0010
0304	8	1.	1.0	7.9	-1.	30.	30.	75.	.32	.28	.29	2.4	5.0	1.00	.72	13.	15.3	1.17	17.9	4.	0.	24.	.6	6.2	0011
0306	9	1.	.7	7.9	-1.	55.	55.	75.	.23	.29	.31	.5	9.0	1.62	.77	13.	14.8	1.09	16.1	-1.	-1.	-1.	-1.	-1.	1000
0307	9	1.	.7	7.9	-1.	51.	55.	75.	.24	.25	.26	1.0	3.0	.54	.34	9.	15.4	.82	12.6	-1.	-1.	-1.	-1.	-1.	0000
5																									
0401	9	1.	.7	7.5	-1.	79.	79.	-1.	.13	.10	.10	2.5	3.5	.63	.53	4.	5.5	1.09	6.0	-1.	-1.	-1.	-1.	-1.	0000
0402	9	0.	.0	7.5	-1.	65.	65.	-1.	.15	.21	.21	.0	4.5	.82	.37	8.	13.2	1.09	14.4	-1.	-1.	-1.	-1.	-1.	0000
0403	9	0.	.0	7.5	-1.	58.	58.	-1.	.17	.21	.20	.5	3.5	.63	.33	8.	13.9	1.09	15.2	-1.	-1.	-1.	-1.	-1.	0000
0404	9	1.	.7	7.5	-1.	37.	37.	-1.	.24	.28	.28	1.0	8.0	1.44	.74	3.	3.5	1.09	3.8	-1.	-1.	-1.	-1.	-1.	0100
0405	9	1.	.7	7.5	-1.	13.	13.	-1.	.40	.47	.47	1.0	7.5	.82	.70	4.	4.8	1.09	5.2	-1.	-1.	-1.	-1.	-1.	0000
8																									
0502	4	1.	1.0	8.3	25.	48.	16.	-1.	.31	.30	.34	2.4	7.6	1.36	.85	16.	17.4	.75	13.1	-1.	-1.	-1.	.28	1.5	0000
0504	6	1.	1.0	8.3	25.	48.	16.	-1.	.31	.31	.34	1.5	4.6	.83	.52	7.	9.7	.75	7.3	-1.	-1.	-1.	-1.	-1.	0000
0505	6	1.	1.0	8.3	25.	48.	16.	-1.	.31	.25	.28	3.0	5.0	.90	.69	5.	6.0	.75	4.5	-1.	-1.	-1.	-1.	-1.	0000
0506	6	1.	1.0	8.3	25.	184.	30.	-1.	.20	.25	.28	0.5	5.5	.99	.49	8.5	12.2	.75	9.2	-1.	-1.	-1.	-1.	-1.	0000
0507	9	1.	1.0	8.3	25.	48.	16.	-1.	.31	.30	.34	1.7	4.6	.83	.53	8.	11.0	1.05	11.6	-1.	-1.	-1.	.28	1.5	0000
0508	9	0.	.0	8.3	25.	48.	16.	-1.	.31	.30	.34	1.7	4.6	.83	.53	14.	19.2	1.05	20.2	-1.	-1.	-1.	.28	1.5	0000
2																									
0601	9	0.	.0	6.9	-1.	51.	51.	-1.	.13	.17	.15	.5	9.0	1.62	.77	13.	14.8	1.09	16.1	-1.	-1.	-1.	-1.	-1.	1000
0602	9	0.	.0	6.9	-1.	57.	57.	-1.	.12	.12	.11	1.0	3.0	.54	.34	9.	15.4	.82	12.6	-1.	-1.	-1.	-1.	-1.	0000
12																									
0702	7	1.	.7	7.9	-1.	72.	72.	-1.	.20	.17	.18	4.0	8.0	1.44	1.04	1.	1.0	1.09	1.1	14.	15.	0.	.25	3.4	0100
0703	8	1.	.7	7.9	-1.	72.	72.	-1.	.20	.13	.14	4.0	4.3	.76	.73	2.	2.3	1.09	2.6	22.	11.	3.	.18	45.0	1101
0704	8	1.	1.0	7.9	-1.	72.	72.	-1.	.20	.23	.24	1.0	8.0	1.58	.85	16.	17.4	1.21	21.0	1.	0.	0.	.28	1.6	0000
0705	8	1.	1.0	7.9	-1.	72.	72.	-1.	.20	.22	.23	1.0	5.0	1.00	.57	12.	15.9	1.09	17.3	5.	0.	0.	.30	2.7	0000
0707	7	0.	.5	7.9	-1.	72.	72.	-1.	.20	.18	.19	3.0	8.0	1.52	1.02	2.	2.0	1.09	2.2	33.	9.	0.	.15	25.0	1101
0708	9	0.	.0	7.9	-1.	114.	114.	-1.	.12	.09	.09	2.5	3.5	.63	.53	4.	5.5	1.09	6.0	-1.	-1.	-1.	-1.	-1.	0000
0709	9	0.	.0	7.9	-1.	102.	102.	-1.	.13	.18	.19	.0	4.5	.82	.37	8.	13.2	1.09	14.4	-1.	-1.	-1.	-1.	-1.	0000
0710	9	0.	.0	7.9	-1.	105.	105.	-1.	.13	.16	.17	.5	3.5	.63	.33	8.	13.9	1.09	15.2	-1.	-1.	-1.	-1.	-1.	0000
0711	9	1.	.7	7.9	-1.	98.	98.	-1.	.14	.17	.18	1.0	8.0	1.44	.74	3.	3.5	1.09	3.8	-1.	-1.	-1.	-1.	-1.	0100
0712	9	1.	.7	7.9	-1.	77.	77.	-1.	.17	.20	.21	1.0	7.5	1.35	.70	4.	4.8	1.09	5.2	-1.	-1.	-1.	-1.	-1.	0000
0713	9	1.	.7	7.9	-1.	72.	72.	-1.	.18	.13	.14	1.9	2.5	.48	.42	4.	6.2	.82	5.1	-1.	-1.	-1.	-1.	-1.	1000
0714	9	1.	.7	7.9	-1.	75.	75.	-1.	.17	.14	.15	1.3	2.2	.40	.31	6.	10.7	.82	8.8	-1.	-1.	-1.	-1.	-1.	1000
1																									
0801	2	1.	1.0	6.3	-1.	11.	11.	15.	.20	.18	.13	4.6	7.6	1.36	1.06	3.	2.8	.75	2.1	-1.	-1.	-1.	-1.	-1.	0000
5																									
0901	9	1.	.7	7.0	-1.	20.	20.	-1.	.25	.19	.17	2.5	3.5	.63	.53	4.	5.5	1.09	6.0	-1.	-1.	-1.	-1.	-1.	0000
0902	9	1.	.7	7.0	-1.	19.	19.	-1.	.25	.35	.32	.0	4.5	.82	.37	8.	13.2	1.09	14.4	-1.	-1.	-1.	-1.	-1.	0000
0903	9	0.	.0	7.0	-1.	34.	34.	-1.	.18	.22	.20	.5	3.5	.63	.33	8.	13.9	1.09	15.2	-1.	-1.	-1.	-1.	-1.	0000
0904	9	1.	.7	7.0	-1.	51.	51.	-1.	.14	.17	.15	1.0	8.0	1.44	.74	3.	3.5	1.09	3.8	-1.	-1.	-1.	-1.	-1.	0100
0905	9	0.	.0	7.0	-1.	65.	65.	-1.	.11	.13	.12	1.0	7.5	1.35	.70	4.	4.8	1.09	5.2	-1.	-1.	-1.	-1.	-1.	0000
5																									
1001	9	0.	.0	6.3	16.	24.	8.	-1.	.21	.20	.18	3.0	7.9	1.43	.94	10.	10.3	.75	7.7	13.	0.	0.	.1	1.5	1100
1002	9	0.	.0	6.3	16.	24.	8.	-1.	.21	.17	.14	5.5	11.0	1.98	1.43	8.	6.7	.75	5.0	13.	0.	0.	.1	1.5	1100
1003	9	0.	.0	6.3	16.	24.	8.	-1.	.21	.19	.16	3.0	7.3	1.32	.89	7.	7.4	.75	5.6	13.	0.	0.	.1	1.5	1100
1004	9	0.	.0	6.3	16.	24.	8.	-1.	.21	.20	.18	3.0	6.4	.95	.61	13.	16.7	.75	12.5	13.	0.	0.	.1	1.5	1100
1005	8	0.	.0	6.3	16.	24.	8.	-1.	.21	.21	.17	1.8	8.2	1.51	.90	8.	8.4	.75	6.3	25.	0.	0.	.11	-1.	1100
3																									
1101	2	1.	1.0	7.0	-1.	8.	8.	-1.	.25	.16	.14	4.6	4.6	.77	.77	9.	10.2	.75	7.7	-1.	-1.	-1.	-1.	-1.	0000
1102	2	1.	1.0	7.0	-1.	8.	8.	-1.	.25	.17	.16	6.1	7.6	1.32	1.17	4.	3.7	.75	2.8	-1.	-1.	-1.	-1.	-1.	0000
1103	2	1.	1.0	7.0	-1.	8.	8.	-1.	.25	.26	.24	1.5	6.1	1.15	.69	1.	1.2	.75	.9	-1.	-1.	-1.	-1.	-1.	0000
4																									
1201	9	1.	1.0	8.0	-1.	165.	60.	70.	.20	.14	.15	2.0	2.4	.43	.39	1.	1.6	1.17	1.9	25.	10.	5.	.3	30.0	0101
1202	8	1.	1.0	8.0	-1.	165.	60.	70.	.20	.22	.24	.5	3.7	.70	.40	1.	1.6	1.17	1.9	27.	12.	2.	.2	67.0	1101

Table A.2 (Continued)

CASE	ECAT	LIQ1	LIQ2	M	K	EP	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC	GRAD
1203	8	1.	1.0	8.0	-1.	165.	80.	70.	.20	.14	.15	2.5	3.0	.55	.49	2.	2.9	1.17	3.4	30.	4.	0.	.15	7.5	1101
1204	8	1.	1.0	8.0	-1.	200.	100.	70.	.16	.15	.17	2.0	7.0	1.35	.86	10.	10.8	1.17	12.6	5.	0.	0.	.28	2.2	0000
1																									
1304	2	1.	1.0	7.3	-1.	6.	6.	30.	.40	.37	.36	1.2	3.0	.56	.38	3.	4.9	.88	4.3	0.	0.	0.	.45	1.4	0000
1																									
1401	3	0.	.0	5.4	0.	40.	40.	-1.	.04	.04	.03	2.4	17.1	2.97	1.50	22.	17.9	.75	13.4	-1.	-1.	-1.	-1.	-1.	0000
32																									
1501	3	0.	.0	5.3	9.	11.	11.	18.	.14	.10	.07	4.6	6.1	1.09	.94	4.	4.1	.75	3.1	-1.	-1.	-1.	-1.	-1.	0000
1502	8	1.	1.0	5.3	9.	6.	6.	18.	.18	.13	.09	2.4	3.0	.52	.46	5.	7.4	.75	5.5	3.	0.	0.	.34	1.9	0000
1503	3	0.	.0	5.3	9.	10.	10.	18.	.15	.10	.07	3.7	4.0	.71	.68	14.	16.9	.75	12.7	-1.	-1.	-1.	-1.	-1.	0000
1505	3	0.	.0	5.3	9.	16.	16.	18.	.10	.10	.07	2.4	7.6	1.37	.85	16.	17.4	.75	13.1	-1.	-1.	-1.	.28	1.5	0000
1507	3	0.	.0	5.3	9.	13.	13.	18.	.12	.12	.08	1.5	4.6	.83	.52	24.	33.3	.75	25.0	-1.	-1.	-1.	-1.	-1.	0000
1508	3	0.	.0	5.3	9.	11.	11.	18.	.14	.15	.10	1.5	6.1	1.09	.63	6.	7.5	.75	5.6	-1.	-1.	-1.	-1.	-1.	0000
1509	3	0.	.0	5.3	9.	16.	16.	18.	.10	.07	.05	4.6	6.1	1.09	.94	20.	20.6	.75	15.5	-1.	-1.	-1.	-1.	-1.	0000
1510	3	0.	.0	5.3	9.	16.	16.	18.	.10	.09	.06	2.4	4.6	.82	.60	20.	25.7	.56	19.3	-1.	-1.	-1.	-1.	-1.	0000
1511	3	0.	.0	5.3	9.	11.	11.	18.	.14	.10	.07	.9	1.2	.22	.19	4.	9.2	.56	5.2	-1.	-1.	-1.	-1.	-1.	0000
1512	3	0.	.0	5.3	9.	11.	11.	18.	.14	.10	.07	.9	1.2	.22	.19	6.	13.9	.75	7.8	-1.	-1.	-1.	-1.	-1.	0000
1514	3	0.	.0	5.3	9.	11.	11.	18.	.14	.15	.10	1.2	4.3	.77	.46	6.	8.8	.75	6.6	-1.	-1.	-1.	-1.	-1.	0000
1515	3	0.	.0	5.3	9.	16.	16.	18.	.10	.07	.05	3.1	3.7	.66	.60	11.	14.2	.75	10.7	-1.	-1.	-1.	-1.	-1.	0000
1516	3	0.	.0	5.3	9.	16.	16.	18.	.10	.07	.05	3.1	4.0	.71	.62	10.	12.7	.75	9.5	-1.	-1.	-1.	-1.	-1.	0000
1517	3	0.	.0	5.3	9.	24.	24.	18.	.07	.07	.05	1.8	5.8	1.04	.64	12.	15.0	.75	11.3	-1.	-1.	-1.	-1.	-1.	0000
1518	3	0.	.0	5.3	9.	24.	24.	18.	.07	.08	.05	1.2	5.8	1.04	.58	16.	20.9	.75	15.7	-1.	-1.	-1.	-1.	-1.	0000
1519	3	0.	.0	5.3	9.	24.	24.	18.	.07	.06	.04	1.8	3.7	.67	.48	10.	14.5	.75	10.9	-1.	-1.	-1.	-1.	-1.	0000
1520	3	0.	.0	5.3	9.	20.	20.	18.	.08	.08	.05	2.4	7.6	1.37	.85	3.	3.3	.75	2.5	-1.	-1.	-1.	-1.	-1.	0000
1521	3	0.	.0	5.3	9.	20.	20.	18.	.08	.08	.05	2.4	9.1	1.64	.97	5.	5.1	.75	3.8	-1.	-1.	-1.	-1.	-1.	0000
1522	3	0.	.0	5.3	9.	20.	20.	18.	.08	.08	.05	1.8	5.8	1.04	.64	7.	8.7	.75	6.5	-1.	-1.	-1.	-1.	-1.	0000
1523	3	0.	.0	5.3	9.	20.	20.	18.	.08	.08	.05	1.8	5.8	1.04	.64	5.	6.2	.75	4.7	-1.	-1.	-1.	-1.	-1.	0000
1525	3	0.	.0	5.3	9.	20.	20.	18.	.08	.08	.05	1.8	4.6	.83	.55	5.	6.8	.75	5.1	-1.	-1.	-1.	-1.	-1.	0000
1526	3	0.	.0	5.3	9.	20.	20.	18.	.08	.08	.05	1.8	4.6	.78	.50	8.	11.3	.75	8.5	-1.	-1.	-1.	-1.	-1.	0000
1527	3	0.	.0	5.3	9.	20.	20.	18.	.08	.08	.05	1.8	4.6	.78	.50	5.	7.0	.75	5.3	-1.	-1.	-1.	-1.	-1.	0000
1528	3	0.	.0	5.3	9.	20.	20.	18.	.08	.08	.05	1.8	4.6	.78	.50	15.	21.3	.75	16.0	-1.	-1.	-1.	-1.	-1.	0000
1529	3	0.	.0	5.3	9.	22.	22.	18.	.08	.07	.04	1.8	3.1	.56	.43	3.	4.6	.75	3.5	-1.	-1.	-1.	-1.	-1.	0000
1530	3	0.	.0	5.3	9.	22.	22.	18.	.08	.06	.04	2.1	2.7	.49	.43	7.	10.6	.56	5.9	-1.	-1.	-1.	-1.	-1.	0000
1531	3	0.	.0	5.3	9.	22.	22.	18.	.08	.08	.05	.6	1.5	.27	.18	13.	30.6	.56	17.1	-1.	-1.	-1.	-1.	-1.	0000
1532	3	0.	.0	5.3	9.	26.	26.	18.	.06	.06	.04	1.2	4.3	.77	.46	7.	10.3	.75	7.7	-1.	-1.	-1.	-1.	-1.	0000
1533	3	0.	.0	5.3	9.	20.	20.	18.	.08	.09	.06	1.2	4.6	.75	.41	5.	7.8	.75	5.9	-1.	-1.	-1.	-1.	-1.	0000
1534	3	0.	.0	5.3	9.	20.	20.	18.	.08	.09	.06	1.2	3.7	.60	.35	12.	20.2	.75	15.2	-1.	-1.	-1.	-1.	-1.	0000
1535	9	0.	.0	5.3	9.	16.	16.	18.	.10	.10	.07	1.7	4.6	.83	.53	8.	11.0	1.05	11.6	-1.	-1.	-1.	.28	1.5	0000
1536	9	0.	.0	5.3	9.	16.	16.	18.	.10	.10	.07	1.7	4.6	.83	.53	14.	19.2	1.05	20.2	-1.	-1.	-1.	.28	1.5	0000
6																									
1601	2	1.	1.0	8.4	50.	340.	113.	75.	.15	.11	.12	3.7	4.6	.84	.75	6.	6.9	.75	5.2	-1.	-1.	-1.	-1.	-1.	0000
1602	2	1.	1.0	8.4	50.	340.	113.	75.	.15	.11	.12	3.7	4.6	.84	.75	8.	9.2	.75	6.2	-1.	-1.	-1.	-1.	-1.	0000
1603	2	0.	.0	8.4	50.	340.	113.	75.	.15	.12	.13	3.7	6.1	1.14	.90	15.	15.8	.75	11.9	-1.	-1.	-1.	-1.	-1.	0000
1604	1	0.	.0	7.5	50.	45.	15.	75.	.15	.12	.12	3.7	7.0	1.29	.97	10.	10.2	.75	7.7	-1.	-1.	-1.	-1.	-1.	0000
1605	1	0.	.0	7.5	50.	45.	15.	75.	.15	.13	.13	3.7	7.9	1.46	1.04	35.	34.3	.75	25.7	-1.	-1.	-1.	-1.	-1.	0000
5																									
1701	2	1.	1.0	8.3	20.	142.	97.	180.	.15	.20	.22	.0	6.1	1.16	.55	7.	9.4	.75	7.1	18.	0.	11.	.25	-1.	0111
1702	2	1.	1.0	8.3	20.	142.	97.	180.	.15	.14	.15	2.4	6.1	1.12	.75	5.	5.8	.75	4.4	18.	0.	11.	.25	-1.	0111
1703	2	0.	.0	8.3	20.	145.	113.	180.	.12	.15	.16	.0	7.6	1.51	.75	42.	48.5	.75	36.4	7.	0.	55.	6.0	40.0	0011
1704	2	1.	1.0	8.3	20.	126.	89.	180.	.16	.21	.23	.0	6.1	1.16	.55	10.	13.5	.75	10.1	10.	0.	6.	.15	4.7	0000
1705	2	1.	1.0	8.3	20.	69.	56.	180.	.25	.26	.29	1.5	6.1	1.11	.65	13.	16.1	.75	12.1	6.	0.	35.	1.8	16.0	0011
11																									
1804	2	0.	.0	7.5	35.	51.	51.	40.	.16	.15	.15	3.7	7.6	1.14	.75	8.	9.2	1.09	10.0	2.	0.	0.	.3	2.5	0000
1814	8	1.	1.0	7.5	35.	51.	51.	40.	.16	.18	.18	1.0	7.0	1.35	.74	8.	9.3	1.09	10.1	2.	0.	0.	.3	2.5	0000
1815	8	0.	.5	7.5	35.	51.	51.	40.	.16	.18	.18	1.0	7.0	1.35	.74	12.	13.9	1.09	15.2	2.	0.	0.	.3	2.5	0000
1816	8	0.	.0	7.5	35.	51.	51.	40.	.18	.18	.18	1.8	7.0	1.35	.83	18.	19.8	1.21	23.9	2.	0.	0.	.3	2.5	0000
1817	8	1.	1.0	7.5	35.	51.	51.	40.	.16	.18	.18	1.0	10.0	1.90	1.02	10.	9.9	1.09	10.8	2.	0.	0.	.3	2.5	0000
1818	8	0.	.5	7.5	35.	51.	51.	40.	.16	.18	.18	1.0	10.0	1.93	1.02	16.	15.8	1.09	17.3	2.	0.	0.	.3	2.5	0000
1819	8	0.	.0	7.5	35.	51.	51.	40.	.18	.18	.18	2.0	10.0	1.93	1.11	20.	19.0	1.21	23.0	2.	0.	0.	.3	2.5	0000
1820	8	1.	1.0	7.5	35.	51.	51.	40.	.16	.21	.21	.0	4.3	.82	.40	4.	6.3	1.09	6.9	10.	4.	0.	.4	6.0	0000
1821	8	0.	.0	7.5	35.	51.	51.	40.	.18	.19	.19	1.3	6.0	1.17	.68	27.	32.7	1.21	39.6	0.	0.	0.	.3	1.4	0000
1822	8	0.	.0	7.5	35.	51.	51.	40.	.18	.16	.16	2.5	6.0	1.17	.81	12.	13.3	1.09	14.5	0.	0.	0.	.36	1.5	0000
1823	8	1.	1.0	7.5	35.	51.	51.	40.	.16	.18	.18	.6	4.5	.88	.48	6.	8.7	1.09	9.4	0.	0.	0.	.40	1.6	0000
1																									
1901	3	0.	.0	4.9	-1.	24.	24.	-1.	.05	.05	.03	2.4	17.1	2.97	1.50	22.	17.9	.75	13.4	-1.	-1.	-1.	-1.	-1.	0000

Table A.2 (Continued)

CASE	ECAT	LIQ1	LIQ2	M	H	EP	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC	GRAD
1																									
2001	8	1.	1.0	6.3	-1.	56.	56.	15.	.13	.08	.07	.9	.9	.16	.16	3.	7.5	.68	4.2	5.	-1.	-1.	-1.	-1.	0000
7																									
2201	8	1.	1.0	7.9	20.	290.	210.	45.	.20	.21	.22	1.0	4.0	.76	.46	5.	7.4	1.17	8.8	20.	6.	0.	.12	6.0	1101
2206	9	0.	.0	7.9	20.	190.	75.	45.	.23	.23	.24	1.0	2.9	.52	.33	14.	24.2	1.21	38.4	5.	0.	0.	.25	3.0	0000
2207	8	0.	.0	7.9	20.	190.	75.	45.	.23	.21	.23	2.0	6.0	1.17	.78	28.	31.7	1.21	38.4	5.	0.	0.	.25	3.0	0000
2208	8	0.	.0	7.9	20.	190.	75.	45.	.23	.24	.25	1.0	4.0	.76	.46	16.	23.6	1.21	28.5	5.	0.	0.	.25	3.0	0000
2209	9	0.	.0	7.9	20.	190.	75.	45.	.23	.18	.19	1.6	2.5	.45	.36	11.	18.3	.88	16.1	5.	0.	0.	.25	3.0	0000
2210	9	1.	1.0	7.9	20.	190.	75.	45.	.23	.19	.20	1.2	2.0	.36	.28	1.	1.9	.88	1.7	5.	0.	0.	.25	3.0	0000
2212	9	0.	.0	7.9	20.	190.	75.	45.	.24	.21	.22	1.3	2.5	.45	.33	5.	8.7	.88	7.7	-1.	-1.	-1.	-1.	-1.	0100
6																									
2301	9	0.	.0	6.1	50.	48.	48.	-1.	.08	.05	.04	8.0	10.0	1.80	1.60	14.	11.1	1.09	12.1	-1.	-1.	-1.	-1.	-1.	0100
2302	4	0.	.0	6.1	50.	48.	48.	-1.	.08	.05	.04	8.0	10.0	1.80	1.60	47.	37.1	1.21	44.9	-1.	-1.	-1.	-1.	-1.	1000
2303	4	0.	.0	6.1	50.	48.	48.	-1.	.08	.08	.06	2.0	6.3	1.13	.70	10.	12.0	1.09	13.1	-1.	-1.	-1.	-1.	-1.	1000
2304	9	0.	.0	6.1	50.	48.	48.	-1.	.08	.08	.06	2.0	6.0	1.08	.68	4.	4.9	1.09	5.3	-1.	-1.	-1.	-1.	-1.	1000
2305	9	0.	.0	6.1	50.	48.	48.	-1.	.08	.07	.05	3.5	6.5	1.17	.87	5.	5.4	1.09	5.9	-1.	-1.	-1.	-1.	-1.	0000
2306	9	0.	.0	6.1	50.	61.	61.	-1.	.07	.05	.04	3.0	3.8	.68	.60	5.	6.4	1.09	7.0	-1.	-1.	-1.	-1.	-1.	1000
2																									
2401	9	0.	.0	5.7	10.	80.	80.	12.	.02	.02	.01	1.7	4.6	.83	.53	8.	11.0	1.05	11.6	-1.	-1.	-1.	.28	1.5	0000
2402	9	0.	.0	5.7	10.	80.	80.	12.	.02	.02	.01	1.7	4.6	.83	.53	14.	19.2	1.05	20.2	-1.	-1.	-1.	.28	1.5	0000
1																									
2501	4	0.	.0	7.1	-1.	130.	130.	-1.	.07	.06	.05	3.7	7.0	1.26	.93	12.	12.4	.75	9.3	-1.	-1.	-1.	-1.	-1.	0000
3																									
2601	8	1.	1.0	6.6	11.	14.	8.	15.	.45	.33	.28	4.6	6.1	1.15	.98	2.	2.0	.75	1.5	50.	-1.	-1.	-1.	-1.	1100
2602	2	1.	1.0	6.6	11.	14.	8.	15.	.45	.21	.18	16.8	16.8	3.30	3.30	24.	13.2	.75	9.9	50.	-1.	0.	.07	-1.	1100
2603	8	1.	1.0	6.6	11.	14.	8.	15.	.45	.35	.30	3.0	6.0	1.10	.88	9.	9.6	.75	7.2	20.	-1.	0.	.1	-1.	1100
1																									
2701	4	0.	.0	7.3	50.	280.	280.	-1.	.01	.01	.01	3.1	16.0	2.88	1.59	10.	7.9	1.09	8.6	-1.	-1.	-1.	-1.	-1.	0100
11																									
2803	5	0.	.0	7.3	12.	90.	90.	-1.	.10	.10	.09	2.0	6.0	1.14	.73	8.5	11.1	1.00	11.1	-1.	-1.	-1.	-1.	-1.	1000
2805	5	0.	.0	7.3	12.	104.	104.	-1.	.10	.09	.09	2.0	13.0	2.58	1.48	14.5	11.9	1.00	11.9	-1.	-1.	-1.	-1.	-1.	0000
2806	5	1.	1.0	7.3	12.	95.	95.	-1.	.10	.10	.09	1.5	5.2	1.02	.65	5.5	6.8	1.00	6.8	-1.	-1.	-1.	-1.	-1.	1100
2808	5	1.	1.0	7.3	12.	57.	57.	-1.	.13	.14	.14	1.5	7.0	1.26	.71	6.	7.1	1.00	7.1	-1.	-1.	-1.	-1.	-1.	1100
2809	8	1.	1.0	7.3	12.	53.	53.	-1.	.20	.21	.20	1.5	8.2	1.61	.94	11.	11.3	1.00	11.3	-1.	-1.	-1.	-1.	-1.	1100
2810	5	1.	1.0	7.3	12.	60.	60.	-1.	.13	.10	.10	2.0	3.0	.58	.48	6.	8.7	.75	8.5	-1.	-1.	-1.	-1.	-1.	0000
2811	5	1.	1.0	7.3	12.	55.	55.	-1.	.20	.20	.19	2.0	10.0	1.98	1.18	9.	8.3	1.00	8.3	-1.	-1.	-1.	-1.	-1.	0000
2812	5	1.	1.0	7.3	12.	55.	55.	-1.	.20	.21	.21	2.0	10.3	1.85	1.02	9.	8.9	1.00	8.9	-1.	-1.	-1.	-1.	-1.	0000
2813	8	1.	1.0	7.3	12.	55.	55.	-1.	.13	.14	.13	1.5	9.1	1.79	1.03	8.	8.0	.85	8.6	67.	-1.	-1.	.064	-1.	1100
2814	8	1.	1.0	7.3	12.	50.	50.	-1.	.20	.21	.20	1.5	8.2	1.61	.94	13.	13.5	1.00	13.5	48.	-1.	-1.	.078	-1.	1100
2815	8	0.	.0	7.3	12.	90.	90.	-1.	.10	.10	.10	1.5	8.2	1.61	.94	9.	9.0	1.00	9.0	-1.	-1.	-1.	-1.	-1.	1100
4																									
2901	9	1.	1.0	7.5	-1.	170.	40.	-1.	.13	.16	.16	1.5	8.8	1.41	.68	3.	3.6	.75	2.7	5.	0.	5.	1.0	6.0	0001
2902	9	0.	.5	7.5	-1.	170.	40.	-1.	.13	.15	.15	2.6	5.8	.72	.40	7.	11.0	.75	8.3	5.	0.	5.	1.0	6.0	0001
2903	9	0.	.0	7.5	-1.	170.	40.	-1.	.13	.14	.14	4.1	16.8	2.21	.94	12.	12.3	.75	9.3	5.	0.	5.	1.0	6.0	0001
2904	9	0.	.0	7.5	-1.	170.	40.	-1.	.13	.15	.15	3.4	17.9	2.42	.97	11.	11.2	.75	8.3	5.	0.	5.	1.0	6.0	0001
8																									
3001	5	1.	1.0	7.8	14.	43.	43.	-1.	.20	.17	.17	1.4	2.3	.45	.33	11.	19.1	.75	14.4	-1.	-1.	-1.	-1.	-1.	1000
3002	5	1.	1.0	7.8	14.	25.	25.	-1.	.35	.39	.41	1.0	7.0	1.34	.74	4.	4.6	1.00	4.6	-1.	-1.	-1.	-1.	-1.	1000
3006	8	0.	.0	7.6	14.	1.	1.	-1.	.50	.40	.41	3.0	5.3	1.00	.77	30.	33.5	1.00	33.5	10.	0.	0.	.196	3.1	1000
3007	8	1.	1.0	7.6	14.	20.	20.	-1.	.35	.38	.38	.9	5.3	1.04	.60	17.	21.5	1.00	21.5	20.	0.	0.	.137	2.5	1100
3009	8	1.	1.0	7.6	14.	80.	80.	-1.	.20	.15	.15	1.5	2.0	.37	.32	10.	16.5	.75	12.5	12.	0.	0.	.22	3.1	0000
3010	8	1.	1.0	7.6	14.	90.	90.	-1.	.13	.14	.14	1.2	6.1	1.19	.70	10.	11.5	1.00	11.5	12.	0.	0.	.14	2.1	1000
3011	9	1.	1.0	7.6	14.	90.	90.	-1.	.20	.21	.21	.9	4.2	.83	.50	6.	8.5	1.00	8.5	-1.	-1.	-1.	-1.	-1.	1100
3012	9	0.	.0	7.6	14.	110.	110.	-1.	.07	.06	.06	3.9	10.2	1.94	1.31	3.	2.8	1.00	2.8	1.	0.	0.	.30	1.7	0000
3013	9	1.	.0	7.6	14.	110.	110.	-1.	.20	.20	.20	1.5	6.2	1.18	.71	2.	2.4	1.00	2.4	2.	0.	0.	.32	2.2	0000
17																									
3201	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.17	.17	4.6	9.9	1.78	1.25	22.	19.6	.75	14.7	10.	0.	0.	.15	-1.	1100
3202	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.17	.17	4.7	11.0	1.98	1.35	8.	6.9	.75	5.2	80.	0.	0.	.08	-1.	1100
3203	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.12	.12	6.9	6.9	1.26	1.26	4.	3.6	.75	2.7	-1.	-1.	-1.	-1.	-1.	0100
3204	9	0.	.0	7.4	40.	70.	70.	55.	.20	.23	.23	1.2	8.1	1.46	.77	7.	8.0	.75	6.0	-1.	-1.	-1.	-1.	-1.	0100
3205	9	0.	.0	7.4	40.	70.	70.	55.	.20	.16	.15	2.1	3.2	.58	.47	18.	26.3	.75	19.7	-1.	-1.	-1.	-1.	-1.	0000
3206	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.20	.19	1.8	5.2	.94	.60	6.	7.7	.75	5.8	-1.	-1.	-1.	-1.	-1.	0000
3207	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.19	.19	3.2	11.0	1.98	1.20	8.	7.3	.75	5.5	31.	-1.	0.	.09	-1.	1100
3208	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.19	.19	3.2	9.0	1.62	1.04	12.	11.8	.75	8.8	20.	-1.	0.	.09	-1.	1100
3209	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.19	.19	3.1	9.0	1.62	1.03	14.	13.8	.75	10.3	11.	-1.	0.	.11	-1.	1000
3210	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.18	.18	3.1	8.0	1.44	.95	12.	12.3	.75	9.2	13.	-1.	0.	.18	-1.	1100
3211	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.19	.19	3.2	10.0	1.80	1.12	8.	7.6	.75	5.7	6.	0.	0.	.12	-1.	1000

Table A.2 (Continued)

CASE	ECAT	LIQ1	LIQ2	M	H	EP	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC	GRAD
3212	9	1.	1.0	7.4	40.	70.	70.	55.	.20	.18	.18	3.4	8.0	1.44	.95	10.	10.3	.75	7.7	7.	0.	0.	.25	3.3	0000
3213	9	0.	.0	7.4	40.	70.	70.	55.	.20	.20	.19	2.7	9.0	1.62	.99	17.	17.0	.75	12.8	5.	0.	41.	3.5	10.0	0011
3214	9	0.	.0	7.4	40.	70.	70.	55.	.20	.19	.19	2.8	12.0	2.16	1.24	10.	9.0	.75	6.7	13.	-1.	10.	.25	-1.	0100
3215	9	0.	.0	7.4	40.	70.	70.	55.	.20	.22	.21	1.6	11.0	1.98	1.04	12.	11.8	.75	8.8	20.	-1.	0.	.12	-1.	1100
3216	9	0.	.0	7.4	40.	70.	70.	55.	.20	.22	.22	1.6	9.0	1.62	.88	6.	6.4	.75	4.8	22.	-1.	0.	.11	-1.	1100
3217 1	9	0.	.0	7.4	40.	70.	70.	55.	.20	.22	.21	1.6	7.0	1.26	.72	13.	15.3	.75	11.5	16.	-1.	0.	.12	-1.	1100
3301 25	7	1.	1.0	7.0	5.	40.	40.	-1.	.25	.28	.28	1.0	7.0	1.33	.73	1.	1.2	1.09	1.3	94.	0.	0.	.03	10.0	1101
3401	8	0.	.0	6.7	60.	140.	140.	-1.	.10	.11	.10	1.0	6.3	1.23	.68	10.	12.1	1.09	13.2	0.	0.	0.	.60	-1.	0000
3402	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.12	.11	.9	3.3	.64	.40	19.	30.0	1.12	33.6	4.	0.	2.	.28	-1.	0000
3403	8	1.	.7	6.7	60.	115.	115.	-1.	.12	.13	.12	.5	3.3	.64	.37	5.	8.2	1.00	8.2	5.	0.	18.	.70	5.0	0010
3404	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.11	.10	1.3	3.3	.64	.43	7.	10.7	1.00	10.7	4.	0.	2.	.28	2.5	0000
3405	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.12	.10	1.8	5.3	1.01	.64	2.	2.5	1.00	2.5	60.	-1.	0.	.04	-1.	1100
3406	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.13	.11	.9	4.3	.82	.48	11.	15.9	1.00	15.9	0.	0.	0.	.4	-1.	0000
3407	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.11	.10	2.2	5.3	1.05	.72	20.	23.6	1.12	28.4	0.	0.	0.	.6	-1.	0000
3408	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.11	.09	1.7	4.3	.82	.58	4.	5.3	1.00	5.3	10.	5.	1.	.4	4.4	0000
3409	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.11	.10	1.8	3.3	.64	.43	13.	19.8	1.12	22.2	7.	0.	43.	1.6	8.0	0011
3410	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.15	.13	.3	4.3	.82	.42	8.	12.3	1.00	12.3	12.	0.	5.	1.2	25.0	0001
3411	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.13	.11	1.3	7.3	1.41	.80	17.	19.0	1.12	21.3	17.	0.	4.	.35	-1.	0100
3412	8	0.	.0	6.7	60.	77.	77.	-1.	.14	.11	.09	4.3	6.3	1.08	.87	9.	9.6	1.00	9.6	5.	0.	0.	.34	2.1	0000
3413	8	0.	.0	6.7	60.	77.	77.	-1.	.14	.13	.12	2.4	6.3	1.17	.77	8.	9.1	1.00	9.1	4.	0.	0.	.36	2.6	0000
3414	8	0.	.0	6.7	60.	77.	77.	-1.	.14	.09	.08	3.0	3.3	.63	.61	11.	14.1	1.00	14.1	5.	0.	0.	.53	2.5	0000
3415	8	0.	.0	6.7	60.	77.	77.	-1.	.14	.12	.10	3.0	6.0	1.17	.87	23.	24.7	1.21	29.8	0.	0.	0.	.41	2.5	0000
3416	8	0.	.0	6.7	60.	75.	75.	-1.	.14	.12	.11	2.5	6.0	1.17	.82	10.	11.0	1.09	12.0	10.	0.	0.	.3	3.5	0000
3417	8	0.	.0	6.7	60.	75.	75.	-1.	.14	.11	.10	2.5	4.0	.76	.61	6.	7.7	1.09	8.4	10.	4.	0.	.25	3.5	0000
3418	8	0.	.0	6.7	60.	75.	75.	-1.	.14	.13	.11	2.5	7.0	1.35	.89	21.	22.3	1.21	28.9	5.	0.	0.	.35	2.5	0000
3419	8	0.	.0	6.7	60.	75.	75.	-1.	.14	.12	.11	2.5	5.0	1.00	.72	9.	10.6	1.09	11.6	20.	7.	0.	.15	15.0	1101
3420	8	0.	.0	6.7	60.	75.	75.	-1.	.14	.12	.10	2.5	4.5	.88	.66	10.	12.3	1.09	13.4	26.	8.	0.	.12	12.0	1101
3421	8	0.	.0	6.7	60.	75.	75.	-1.	.14	.13	.11	2.5	6.0	1.17	.81	12.	13.3	1.09	14.5	3.	0.	2.	.35	2.3	0000
3422	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.11	.10	2.4	6.0	1.13	.77	15.	17.1	1.21	20.7	11.	4.	0.	.30	5.0	0000
3423	8	0.	.0	6.7	60.	115.	115.	-1.	.12	.10	.09	3.6	7.0	1.30	.96	17.	17.4	1.21	21.0	12.	3.	0.	.30	5.0	0000
3424	8	0.	.0	6.7	60.	75.	75.	-1.	.12	.11	.10	1.4	4.0	.76	.52	4.	5.5	1.09	6.0	10.	0.	0.	.15	2.0	1000
3425 25	8	0.	.0	6.7	60.	75.	75.	-1.	.12	.12	.11	1.4	6.0	1.17	.71	15.	17.8	1.21	21.5	10.	0.	0.	.18	2.4	1000
3501	8	1.	1.0	7.4	30.	115.	60.	-1.	.20	.22	.22	1.0	6.3	1.23	.68	10.	12.1	1.09	13.2	0.	0.	0.	.60	-1.	0000
3502	8	0.	.0	7.4	30.	112.	60.	-1.	.32	.32	.32	.9	3.3	.64	.40	19.	30.0	1.12	33.6	4.	0.	2.	.28	-1.	0000
3503	8	1.	1.0	7.4	30.	112.	60.	-1.	.32	.35	.34	.5	3.3	.64	.37	5.	8.2	1.00	8.2	5.	0.	18.	.70	5.0	0010
3504	8	1.	.7	7.4	30.	112.	60.	-1.	.32	.30	.30	1.3	3.3	.64	.43	7.	10.7	1.00	10.7	4.	0.	2.	.28	2.5	0000
3505	8	1.	1.0	7.4	30.	115.	60.	-1.	.24	.24	.23	1.8	5.3	1.01	.64	2.	2.5	1.00	2.5	60.	-1.	0.	.04	-1.	1100
3506	8	1.	.7	7.4	30.	115.	60.	-1.	.24	.26	.25	.9	4.3	.82	.48	11.	15.9	1.00	15.9	0.	0.	0.	.4	-1.	0000
3507	8	0.	.0	7.4	30.	115.	60.	-1.	.24	.22	.21	2.2	5.3	1.05	.72	20.	23.6	1.12	26.4	0.	0.	0.	.6	-1.	0000
3508	8	1.	.7	7.4	30.	115.	60.	-1.	.24	.21	.21	1.7	4.3	.82	.58	4.	5.3	1.00	5.3	10.	5.	1.	.4	4.4	0000
3509	8	1.	1.0	7.4	30.	115.	60.	-1.	.24	.23	.22	1.3	3.3	.64	.43	13.	19.8	1.12	22.2	7.	0.	43.	1.6	8.0	0011
3510	8	1.	.7	7.4	30.	115.	60.	-1.	.24	.29	.29	.3	4.3	.82	.42	8.	12.3	1.00	12.3	12.	0.	5.	1.2	25.0	0001
3511	8	0.	.0	7.4	30.	115.	60.	-1.	.24	.26	.26	1.3	7.3	1.41	.80	17.	19.0	1.12	21.3	17.	0.	4.	.35	-1.	0100
3512	8	1.	.7	7.4	30.	92.	30.	-1.	.24	.18	.18	4.3	6.3	1.08	.87	9.	9.6	1.00	9.6	5.	0.	0.	.34	2.1	0000
3513	8	1.	.7	7.4	30.	92.	30.	-1.	.24	.23	.22	2.4	6.3	1.17	.77	8.	9.1	1.00	9.1	4.	0.	0.	.36	2.6	0000
3514	8	1.	.7	7.4	30.	95.	30.	-1.	.28	.18	.18	3.0	3.3	.63	.61	11.	14.1	1.00	14.1	5.	0.	0.	.53	2.5	0000
3515	8	0.	.0	7.4	30.	95.	30.	-1.	.28	.23	.23	3.0	6.0	1.17	.87	23.	24.7	1.21	29.6	0.	0.	0.	.41	2.5	0000
3516	8	0.	.5	7.4	30.	85.	20.	-1.	.24	.21	.21	2.5	6.0	1.17	.82	10.	11.0	1.09	12.0	10.	0.	0.	.3	3.5	0000
3517	8	1.	.7	7.4	30.	85.	20.	-1.	.24	.19	.19	2.5	4.0	.76	.61	6.	7.7	1.09	8.4	10.	4.	0.	.25	3.5	0000
3518	8	0.	.0	7.4	30.	85.	20.	-1.	.24	.22	.22	2.5	7.0	1.35	.89	21.	22.3	1.21	28.9	5.	0.	0.	.35	2.5	0000
3519	8	1.	.7	7.4	30.	85.	20.	-1.	.24	.21	.20	2.5	5.0	1.00	.72	9.	10.6	1.09	11.6	20.	7.	0.	.15	15.0	1101
3520	8	0.	.5	7.4	30.	85.	20.	-1.	.24	.20	.20	2.5	4.5	.88	.66	10.	12.3	1.09	13.4	26.	8.	0.	.12	12.0	1101
3521	8	1.	1.0	7.4	30.	85.	20.	-1.	.24	.21	.21	2.5	6.0	1.17	.81	12.	13.3	1.09	14.5	3.	0.	2.	.35	2.3	0000
3522	8	0.	.0	7.4	30.	85.	20.	-1.	.24	.22	.21	2.1	6.0	1.13	.77	15.	17.1	1.21	20.7	11.	4.	0.	.30	5.0	0000
3523	8	0.	.0	7.4	30.	85.	20.	-1.	.24	.20	.20	3.6	7.0	1.30	.96	17.	17.4	1.21	21.0	12.	3.	0.	.30	5.0	0000
3524	8	1.	1.0	7.4	30.	80.	20.	-1.	.20	.18	.18	1.4	4.0	.76	.52	4.	5.5	1.09	6.0	10.	0.	0.	.15	2.0	1000
3525 2	8	0.	.0	7.4	30.	80.	20.	-1.	.20	.20	.20	1.4	6.0	1.17	.71	15.	17.8	1.21	21.5	10.	0.	0.	.18	2.4	1000
3601	9	0.	.0	6.5	22.	25.	25.	-1.	.32	.23	.19	4.0	5.0	.90	.80	27.	30.2	.75	22.7	-1.	-1.	-1.	-1.	-1.	0111
3602 2	9	0.	.0	6.5	22.	25.	25.	-1.	.32	.36	.30	1.5	8.7	1.57	.85	7.	7.6	.75	5.7	-1.	-1.	-1.	-1.	-1.	0100
3701	9	0.	.0	7.6	59.	50.	50.	15.	.30	.23	.24	3.5	5.5	1.00	.80	20.	22.4	.75	16.8	8.	0.	3.	.22	3.3	1000
3702	9	1.	1.0	7.6	59.	50.	50.	15.	.30	.26	.27	1.2	2.7	.55	.40	16.	25.2	.75	5.7	8.	0.	3.	.22	3.3	1000

Table A.2 (Continued)

CASE	ECAT	LIQ1	LIQ2	M	H	EP	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC	GRAD
1																									
3801	9	1.	1.0	6.9	10.	50.	50.	-1.	.26	.31	.28	1.0	7.0	1.28	.66	6.	7.4	.75	5.6	10.	0.	0.	.3	3.3	0000
24																									
3901	8	0.	.0	6.6	10.	10.	1.6	-1.	.55	.47	.40	1.8	3.7	.69	.51	28.	39.2	1.05	41.2	25.	0.	0.	.11	-1.	1100
3902	8	1.	1.0	6.6	10.	10.	1.6	-1.	.55	.47	.40	1.8	3.7	.69	.51	1.	1.2	1.05	1.3	29.	5.	0.	.12	-1.	1100
3903	8	0.	.0	6.6	10.	10.	1.6	-1.	.55	.49	.42	1.8	4.3	.81	.57	13.	17.2	1.05	18.1	37.	1.	0.	.09	-1.	1100
3904	8	1.	1.0	6.6	10.	42.	3.4	-1.	.24	.28	.23	.2	1.8	.38	.21	3.	6.5	.79	5.2	80.	-1.	-1.	.04	-1.	1100
3905	8	1.	.7	6.6	10.	42.	3.4	-1.	.24	.28	.24	.2	4.3	.85	.45	11.	16.4	1.05	17.2	18.	4.	0.	.15	-1.	1100
3906	9	0.	.0	6.6	10.	50.	15.	-1.	.16	.16	.14	1.2	3.4	.61	.39	2.	4.8	1.05	5.0	-1.	-1.	0.	-1.	-1.	0100
3907	9	0.	.0	6.6	10.	50.	15.	-1.	.16	.16	.14	1.2	3.4	.61	.39	5.	8.0	1.05	8.4	17.	0.	0.	.10	2.0	1100
3908	9	0.	.0	6.6	10.	50.	15.	-1.	.16	.15	.13	1.2	3.1	.56	.37	6.	9.9	1.05	10.4	-1.	-1.	0.	-1.	-1.	0100
3909	9	0.	.0	6.6	10.	50.	15.	-1.	.16	.17	.15	1.2	4.9	.88	.51	4.	5.6	1.05	5.9	-1.	-1.	0.	-1.	-1.	0100
3910	9	0.	.0	6.6	10.	50.	15.	-1.	.16	.17	.14	1.2	4.3	.77	.46	9.	13.3	1.05	14.0	-1.	-1.	0.	-1.	-1.	0100
3911	9	0.	.0	6.6	10.	50.	15.	-1.	.16	.16	.14	1.2	3.8	.68	.42	5.	7.7	1.05	8.1	-1.	-1.	0.	-1.	-1.	1000
3912	9	0.	.0	6.6	10.	55.	15.	-1.	.14	.11	.09	2.7	4.0	.72	.59	15.	19.5	1.05	20.5	13.	1.	0.	.11	2.2	1100
3913	9	0.	.0	6.6	10.	55.	15.	-1.	.14	.11	.09	2.7	4.0	.72	.59	13.	16.9	1.05	17.7	12.	-1.	0.	.11	2.0	1000
3914	9	0.	.0	6.6	10.	60.	20.	-1.	.10	.08	.07	2.5	4.3	.77	.59	3.	3.9	1.05	4.1	91.	12.	0.	.039	15.0	1101
3915	9	0.	.0	6.6	10.	60.	20.	-1.	.10	.07	.06	2.5	3.0	.54	.49	3.	4.3	.79	3.4	62.	9.	0.	.056	9.0	1101
3916	9	0.	.0	6.6	10.	60.	20.	-1.	.10	.08	.07	2.5	3.7	.67	.55	7.	9.4	1.05	9.9	88.	9.	0.	.042	10.0	1101
3917	9	0.	.0	6.6	10.	60.	20.	-1.	.10	.07	.06	2.5	2.9	.52	.48	1.	1.4	.79	1.1	-1.	-1.	0.	-1.	-1.	1101
3918	9	1.	.7	6.6	10.	45.	5.	-1.	.15	.12	.10	2.1	3.4	.62	.50	2.	2.8	1.05	3.0	64.	5.	0.	.047	4.0	1100
3919	9	0.	.0	6.6	10.	45.	5.	-1.	.15	.10	.09	2.1	2.3	.41	.39	11.	17.6	.79	13.9	18.	1.	0.	.10	2.6	1100
3920	9	1.	.7	6.6	10.	26.	3.	-1.	.50	.40	.34	1.5	2.3	.41	.33	7.	12.2	.79	9.6	24.	10.	0.	.10	26.0	1101
3921	9	1.	.7	6.6	10.	26.	3.	-1.	.50	.40	.34	1.5	2.3	.41	.33	3.	5.2	.79	4.1	-1.	-1.	-1.	-1.	-1.	0000
3922	9	0.	.0	6.6	10.	53.	15.	-1.	.16	.11	.10	4.1	5.0	.90	.81	15.	16.7	1.05	17.5	-1.	-1.	-1.	-1.	-1.	0000
3923	9	0.	.0	6.6	10.	53.	15.	-1.	.16	.10	.09	4.1	4.6	.83	.81	16.	17.8	1.05	18.7	-1.	-1.	-1.	-1.	-1.	0100
3924	9	0.	.0	6.6	10.	60.	20.	-1.	.10	.08	.07	2.7	4.3	.77	.61	9.	11.5	1.05	12.1	-1.	-1.	-1.	-1.	-1.	1100
1																									
4001	9	1.	.0	6.7	30.	23.	1.	-1.	.58	.59	.52	2.0	5.0	.78	.48	5.	7.2	.75	5.4	-1.	-1.	-1.	-1.	-1.	1100
2																									
4101	8	0.	.0	6.1	20.	15.	15.	-1.	.10	.12	.09	1.0	6.0	1.10	.59	5.	6.5	1.09	7.1	13.	-1.	0.	.18	3.0	1100
4103	8	0.	.0	6.1	20.	15.	15.	-1.	.10	.10	.08	1.0	14.3	2.52	1.25	4.	3.6	1.09	3.9	27.	-1.	-1.	.17	-1.	1100
24																									
4201	8	0.	.0	5.6	7.	50.	50.	-1.	.02	.02	.01	1.8	3.7	.69	.51	28.	39.2	1.05	41.2	25.	0.	0.	.11	-1.	1100
4202	8	0.	.0	5.6	7.	50.	50.	-1.	.02	.02	.01	1.8	3.7	.69	.51	1.	1.2	1.05	1.3	29.	5.	0.	.12	-1.	1100
4203	8	0.	.0	5.6	7.	50.	50.	-1.	.02	.02	.01	1.8	4.3	.81	.57	13.	17.2	1.05	18.1	37.	1.	0.	.09	-1.	1100
4204	8	0.	.0	5.6	7.	15.	13.	-1.	.18	.20	.14	.2	1.8	.36	.21	3.	6.5	.79	5.2	80.	-1.	-1.	.04	-1.	1100
4205	8	0.	.0	5.6	7.	15.	13.	-1.	.18	.21	.15	.2	4.3	.85	.45	11.	16.4	1.05	17.2	18.	4.	0.	.15	-1.	1100
4206	9	1.	.7	5.6	7.	10.	5.	-1.	.26	.28	.18	1.2	3.4	.61	.39	3.	4.8	1.05	5.0	-1.	-1.	0.	-1.	-1.	0100
4207	9	1.	.7	5.6	7.	10.	5.	-1.	.26	.26	.18	1.2	3.4	.61	.39	5.	8.0	1.05	8.4	17.	0.	0.	.10	2.0	1100
4208	9	1.	.7	5.6	7.	10.	5.	-1.	.26	.25	.18	1.2	3.1	.56	.37	6.	9.9	1.05	10.4	-1.	-1.	0.	-1.	-1.	0100
4209	9	1.	.7	5.6	7.	10.	5.	-1.	.26	.28	.20	1.2	4.9	.88	.51	4.	5.6	1.05	5.9	-1.	-1.	0.	-1.	-1.	0100
4210	9	1.	.7	5.6	7.	10.	5.	-1.	.26	.27	.19	1.2	4.3	.77	.46	9.	13.3	1.05	14.0	-1.	-1.	0.	-1.	-1.	0100
4211	9	1.	.7	5.6	7.	10.	5.	-1.	.26	.27	.19	1.2	3.8	.68	.42	5.	7.7	1.05	8.1	-1.	-1.	0.	-1.	-1.	1000
4212	9	0.	.5	5.6	7.	7.	2.	-1.	.26	.20	.14	2.7	4.0	.72	.59	15.	19.5	1.05	20.5	13.	1.	0.	.11	2.2	1100
4213	9	0.	.5	5.6	7.	7.	2.	-1.	.26	.20	.14	2.7	4.0	.72	.59	13.	16.9	1.05	17.7	12.	-1.	0.	.11	2.0	1000
4214	9	1.	.7	5.6	7.	3.	3.	-1.	.28	.23	.16	2.5	4.3	.77	.59	3.	3.9	1.05	4.1	91.	12.	0.	.039	15.0	1101
4215	9	1.	.7	5.6	7.	3.	3.	-1.	.28	.20	.14	2.5	3.0	.54	.49	3.	4.3	.79	3.4	62.	9.	0.	.056	9.0	1101
4216	9	1.	.7	5.6	7.	3.	3.	-1.	.28	.22	.15	2.5	3.7	.67	.55	7.	9.4	1.05	9.9	88.	9.	0.	.042	10.0	1101
4217	9	1.	.7	5.6	7.	3.	3.	-1.	.28	.19	.14	2.5	2.9	.52	.48	1.	1.4	.79	1.1	-1.	-1.	0.	-1.	-1.	1101
4218	9	1.	.7	5.6	7.	14.	12.	-1.	.18	.14	.10	2.1	3.4	.62	.50	2.	2.8	1.05	3.0	64.	5.	0.	.047	4.0	1100
4219	9	0.	.0	5.6	7.	14.	12.	-1.	.18	.12	.09	2.1	2.3	.41	.39	11.	17.6	.79	13.9	18.	1.	0.	.10	2.6	1100
4220	9	0.	.0	5.6	7.	28.	26.	-1.	.12	.10	.07	1.5	2.3	.41	.33	7.	12.2	.79	9.6	24.	10.	0.	.10	26.0	1101
4221	9	0.	.0	5.6	7.	28.	26.	-1.	.12	.10	.07	1.5	2.3	.41	.33	3.	5.2	.79	4.1	-1.	-1.	-1.	-1.	-1.	0000
4222	9	0.	.0	5.6	7.	9.	3.	-1.	.26	.18	.13	4.1	5.0	.90	.81	15.	16.7	1.05	17.5	-1.	-1.	-1.	-1.	-1.	0000
4223	9	0.	.0	5.6	7.	9.	3.	-1.	.26	.17	.12	4.1	4.6	.83	.81	16.	17.8	1.05	18.7	-1.	-1.	-1.	-1.	-1.	0100
4224	9	0.	.0	5.6	7.	4.	2.	-1.	.28	.22	.16	2.7	4.3	.77	.61	9.	11.5	1.05	12.1	-1.	-1.	-1.	-1.	-1.	1100

Table A.3

Previously Published Catalog Identification Codes

EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	GCAS	BCAT	ECAT	REP	TAB
Niigata	1802	Niigata	0101	0	1	1	1	2	8	01111000	01000000
Sado Island	12/1	Niigata	0102	0	2	2	0	2	8	01111000	01000000
Niigata	1877	Niigata	0201	0	1	1	0	2	8	01111000	01000000
Koshigun	7/22	Niigata	0202	0	2	2	0	2	8	01111000	01000000
Mino-Owari	1891	Ogaki	0301	1	0	0	0	2	2	01111011	01010011
	10/28	Ginan	0302	1	14	14	0	2	2	11111111	01010011
		Unuma	0303	1	0	0	0	2	8	11111111	01010011
		Ogase Pond	0304	1	0	0	0	2	8	11111111	01010011
		Ginan West	0305	1	14	14	0	3	0	00100000	00100000
		Saya	0306	1	4	4	0	6	6	00000100	00000100
		Biwajima	0307	1	5	5	0	6	6	00000100	00000100
Tokyo	1894	Tone River	0401	1	8	8	0	6	9	00000100	00000100
	6/20	Gyoda	0402	0	9	9	0	6	9	00000100	00000100
		Kasu	0403	0	10	10	0	6	9	00000100	00000100
		Kasukabe	0404	1	11	11	0	6	9	00000100	00000100
		Ara River	0405	1	12	12	0	6	9	00000100	00000100
San Francisco	1906	Foot of Market-a	0501	1	3	17	0	4	0	00110000	00110000
	4/18	Foot of Market-b	0502	1	3	18	0	4	4	00110000	00110000
		Foot of Market-c	0503	1	3	19	0	3	0	00100000	00100000
		So. of Market	0504	1	20	0	0	6	6	00110100	00110100
		Mission Creek	0505	1	21	0	0	6	6	00000100	00000100
		Salinas	0506	1	0	0	0	6	6	00000100	00000100
		Yerba Buena Cove	0507	1	3	3	0	0	9	00000000	00000000
		Telegraph Hill	0508	0	32	32	0	0	9	00000000	00000000
Gono	1909	Saya	0601	0	4	4	0	6	9	00000100	00000100
	8/14	Biwajima	0602	0	5	5	0	6	9	00000100	00000100
Kanto	1923	Arakawa 7	0701	1	6	6	0	7	0	00000011	00000011
	9/1	Arakawa 7	0702	1	6	6	0	7	7	00000010	00000010
		Arakawa 12	0703	1	6	0	0	7	8	00000011	00000011
		Arakawa 21	0704	1	6	0	0	7	8	00000011	00000011
		Arakawa 30	0705	1	6	0	0	7	8	00000011	00000011
		Arakawa 49	0706	1	6	7	0	7	0	00000011	00000011
		Arakawa 49	0707	0	6	7	0	7	7	00000010	00000010
		Tone River	0708	0	8	8	0	6	9	00000100	00000100
		Gyoda	0709	0	9	9	0	6	9	00000100	00000100
		Kasu	0710	0	10	10	0	6	9	00000100	00000100
		Kasukabe	0711	1	11	11	0	6	9	00000100	00000100
		Ara River	0712	1	12	12	0	6	9	00000100	00000100
		Ukita	0713	1	0	0	0	6	9	00000100	00000100
		Edogawa	0714	1	0	0	0	6	9	00000100	00000100
Santa Barbara	1925	Sheffield Dam	0801	1	0	0	0	2	2	11111100	01001100
	6/29										
Nishi-Saitama	1931	Tone River	0901	1	8	8	0	6	9	00000100	00000100
	9/21	Gyoda	0902	1	9	9	0	6	9	00000100	00000100
		Kasu	0903	0	10	10	0	6	9	00000100	00000100
		Kasukabe	0904	1	11	11	0	6	9	00000100	00000100
		Ara River	0905	0	12	12	0	6	9	00000100	00000100
Long Beach	1933	LNG Ter./Res. Pt-1	1001	0	13	0	13	3	9	00100000	00100000
	3/10	LNG Ter./Res. Pt-2	1002	0	13	0	13	3	9	00100000	00100000
		LNG Ter./Res. Pt-3	1003	0	13	0	13	3	9	00100000	00100000
		LNG Ter./Res. Pt-4	1004	0	13	0	13	3	9	00100000	00100000
		L.A. Pier A	1005	0	0	0	0	0	8	00100001	00100001
		Reservation P.	1006	0	13	0	13	8	0	00000001	00000001

Table A.3 (Continued)

EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	GCAS	BCAT	ECAT	REP	TAB
El Centro	1940	Brawley	1101	1	0	0	0	2	2	01001100	01001100
		All-Am. Canal	1102	1	0	0	0	2	2	01001100	01001100
		Solfatara Canal	1103	1	0	0	0	2	2	01001100	01001100
Tonankai	1944	Komei	1201	1	0	0	0	2	9	01111111	01010011
	12/7	Meiko St.	1202	1	0	0	0	2	8	01111111	01010011
		Isnaga	1203	1	0	0	0	7	8	00010011	00010011
		Ginan	1204	1	14	14	0	7	8	00000011	00000011
Fukui	1948	Takaya 2	1301	0	15	15	0	7	7	00000010	00000010
	6/28	Takaya 2	1302	0	15	15	0	2	0	01110111	01010011
		Takaya 45	1303	1	0	0	0	2	8	01110111	01010011
		Shonenji	1304	1	0	0	0	2	2	01110111	01010011
		Agr. Union	1305	0	16	16	0	2	2	01110110	01010010
		Agr. Union	1306	0	16	16	0	2	0	00000011	00000011
San Francisco Concord Bay	1955	Joaquin Aqueduct	1401	0	29	0	0	3	3	00110000	00100000
San Francisco	1957	St. Francis Circle	1501	0	0	0	0	3	3	00110000	00100000
Daly City	3/22	Lake Merced	1502	1	0	0	0	2	8	01111111	01000011
		Duboce & Sanchez	1503	0	0	0	0	3	3	00110000	00100000
		Foot of Market-a	1504	0	3	17	0	3	3	00110000	00100000
		Foot of Market-b	1505	0	3	18	0	3	3	00110000	00100000
		Foot of Market-c	1506	0	3	19	0	3	0	00110000	00100000
		So. of Market	1507	0	20	0	0	3	3	00110000	00100000
		Mission Creek	1508	0	21	0	0	3	0	00110000	00100000
		Polk & Golden Gate	1509	0	0	0	0	3	3	00110000	00100000
		Polk & Market	1510	0	0	0	0	3	3	00110000	00100000
		Welden-a	1511	0	22	0	0	3	3	00110000	00100000
		Welden-b	1512	0	22	0	0	3	3	00110000	00100000
		Welden-c	1513	0	22	0	0	3	0	00110000	00100000
		Welden-d	1514	0	22	0	0	3	3	00110000	00100000
		Mission & Spear-a	1515	0	23	0	0	3	3	00110000	00100000
		Mission & Spear-b	1516	0	23	0	0	3	3	00110000	00100000
		Park & Otis,Al.-a	1517	0	24	0	0	3	3	00110000	00100000
		Park & Otis,Al.-b	1518	0	24	0	0	3	3	00110000	00100000
		Singleton,Alameda	1519	0	0	0	0	3	3	00110000	00100000
		Treasure Island-a	1520	0	25	0	0	3	3	00110000	00100000
		Treasure Island-b	1521	0	25	0	0	3	3	00110000	00100000
		Treasure Island-c	1522	0	25	0	0	3	3	00110000	00100000
		Treasure Island-d	1523	0	25	0	0	3	3	00110000	00100000
		Treasure Island-e	1524	0	25	0	0	3	0	00110000	00100000
		Treasure Island-f	1525	0	25	0	0	3	3	00110000	00100000
		Treasure Island-g	1526	0	25	0	0	3	3	00110000	00100000
		Treasure Island-h	1527	0	25	0	0	3	3	00110000	00100000
		Treasure Island-i	1528	0	25	0	0	3	3	00110000	00100000
		W 5th/Ave D,Al.-a	1529	0	26	0	0	3	3	00110000	00100000
		W 5th/Ave D,Al.-b	1530	0	26	0	0	3	3	00110000	00100000
		Westline Ave,Al.	1531	0	0	0	0	3	3	00110000	00100000
		Emeryville,Al.	1532	0	0	0	0	3	3	00110000	00100000
		Westline M.C.-a	1533	0	27	0	0	3	3	00110000	00100000
		Westline M.C.-b	1534	0	27	0	0	3	3	00110000	00100000
		Yerba Buena Cove	1535	0	3	3	0	0	9	00000000	00000000
		Telegraph Hill	1536	0	32	32	0	0	9	00000000	00000000
Chile	1960	Puerto Montt-a	1601	1	64	0	0	2	2	01111000	01000000
	5/22	Puerto Montt-b	1602	1	64	0	0	2	2	01111000	01000000
		Puerto Montt-c	1603	0	0	0	0	2	2	01111000	01000000
		Concepcion	1604	0	0	0	0	1	1	01111000	01001000
		Huachipato	1605	0	0	0	0	1	1	01111000	01001000
Alaska	1964	Snow River B605A	1701	1	65	65	0	2	2	01111110	01010010
	3/27	Snow River B605	1702	1	65	65	0	2	2	01111110	01010010
		Quartz Creek	1703	0	0	0	0	2	2	01111110	01010010
		Scott Glacier	1704	1	0	0	0	2	2	01111110	01010010

Table A.3 (Continued)

EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	GCAS	BCAT	ECAT	REP	TAB
		Valdez	1705	1	0	0	0	2	2	01111110	01010010
		River Bridges	1706	1	55	55	0	1	0	10000000	10000000
Niigata	1964	Niigata	1801	1	1	0	0	2	0	01101100	01001000
	6/16	Niigata	1802	1	1	0	0	2	0	01101000	01001000
		Niigata	1803	0	1	0	0	2	0	01101100	01001000
		Niigata	1804	0	1	0	28	2	2	01101100	01001000
		Niigata-C	1805	1	1	1	1	5	0	10001000	10001000
		Niigata-Fill	1806	0	1	0	28	1	0	10001000	10001000
		Niigata-B	1807	0	1	2	2	1	0	10001000	10001000
		Niigata	1808	0	1	0	0	3	0	00100000	00100000
		Nippon Fire & Mar.	1809	1	1	0	0	4	0	00010000	00010000
		Iribune School	1810	1	1	0	0	4	0	00010000	00010000
		Benton-Cho	1811	1	1	0	0	4	0	00010000	00010000
		Benton-Cho	1812	1	1	0	0	4	0	00010000	00010000
		Ohsaki (1966)	1813	1	1	0	0	4	0	00010000	00010000
		Niigata	1814	1	1	1	1	7	8	00000011	00000011
		Niigata	1815	0	1	2	2	7	8	00000011	00000011
		Niigata	1816	0	1	0	0	7	8	00000011	00000011
		Niigata	1817	1	1	0	0	7	8	00000011	00000011
		Niigata	1818	0	1	0	0	7	8	00000011	00000011
		Niigata	1819	0	1	0	0	7	8	00000011	00000011
		Shova Bridge 2	1820	1	0	0	0	7	8	00000011	00000011
		Shova Bridge 4	1821	0	0	0	0	7	8	00000011	00000011
		Road Site	1822	0	0	0	0	7	8	00000011	00000011
		River Site	1823	1	0	0	0	7	8	00000011	00000011
San Francisco Concord Bay	1965 9/10	Joaquin Aqueduct	1901	0	29	29	0	3	3	00110000	00100000
Caracus	1967	Caraballeda	2001	1	0	0	0	2	8	01001001	01000001
Ebino	1968	Ebino	2101	1	0	0	0	4	0	00010100	00010000
		Yoshimatsu	2102	1	0	0	0	4	0	00010000	00010000
Tokachi-Oki	1968	Nanehama,Hakadote	2201	1	0	0	0	1	8	11111111	11011111
	5/16	Hachinohe-Beach	2202	0	30	0	30	1	0	11101100	11001100
		Hachinohe-Fill	2203	1	30	0	31	1	0	10001000	10001000
		Hachinohe	2204	0	30	0	30	2	0	01101100	01001000
		Hachinohe	2205	1	30	0	31	2	0	01101100	01001000
		Hachinohe P-1	2206	0	30	0	30	4	9	00010000	00010000
		Hachinohe P-2	2207	0	30	0	30	4	8	00010011	00010011
		Hachinohe P-4	2208	0	30	0	30	4	8	00010011	00010011
		Hachinohe P-5	2209	0	30	0	30	4	9	00010000	00010000
		Hachinohe P-6	2210	1	30	0	30	4	9	00010000	00010000
		Hachinohe Plant	2211	1	30	0	31	4	0	00010000	00010000
		Hachinohe Accel.	2212	0	0	0	0	4	9	00010000	00010000
		Hachinohe 3	2213	1	30	0	0	7	0	00000011	00000011
Saitama	1968	Saitama 101-2	2301	0	0	0	0	4	9	00010000	00010000
	7/1	Saitama 105-2	2302	0	0	0	0	4	4	00010000	00010000
		Saitama 119	2303	0	0	0	0	4	4	00010000	00010000
		Saitama 121	2304	0	0	0	0	4	9	00010000	00010000
		Saitama 130	2305	0	0	0	0	4	9	00010000	00010000
		Saitama 602	2306	0	0	0	0	4	9	00010000	00010000
Santa Rosa	1969	Yerba Buena Cove	2401	0	3	3	0	0	9	00000000	00000000
	10/1	Telegraph Hill	2402	0	32	32	0	0	9	00000000	00000000
Gediz,Turkey	1970 3/28	Bursa	2501	0	0	0	0	4	4	00110000	00010000
San Fernando	1971	Juvenile Hall	2601	1	0	0	0	2	8	01001011	01000011
	2/9	Jensen Plant	2602	1	0	0	0	2	2	01110110	01010010
		Van Norman Dam	2603	1	0	0	0	7	8	00000011	00000011
		L.A. Balboa-a	2604	0	33	0	0	3	0	00100000	00100000

Table A.3 (Continued)

EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	GCAS	BCAT	ECAT	REP	TAB
		L.A. Balboa-b	2605	0	33	0	0	3	0	00100000	00100000
		L.A. Balboa-c	2606	0	33	0	0	3	0	00100000	00100000
		L.A. Balboa-d	2607	1	33	0	0	3	0	00100000	00100000
		L.A. Balboa-e	2608	1	33	0	0	3	0	00100000	00100000
		L.A. Balboa-f	2609	1	33	0	0	3	0	00100000	00100000
		L.A. Balboa-g	2610	1	33	0	0	3	0	00100000	00100000
		L.A. Balboa-h	2611	1	33	0	0	3	0	00100000	00100000
Yokohama	1972 12/4	Yokohama	2701	0	0	0	0	4	9	00110000	00010000
Haicheng	1974	Shuangtaihe Gate	2801	0	34	0	34	5	0	00001000	00001000
		Shuangtaihe Gate	2802	0	34	0	34	5	0	00001000	00001000
		Shuang. Eardo Br.	2803	0	35	0	0	5	5	00001001	00001001
		Shuang. Eardo Br.	2804	0	35	0	0	5	0	00001000	00001000
		Shenglitang	2805	0	0	0	0	5	5	00001001	00001001
		Ligohe Plant	2806	1	36	0	0	5	5	00001000	00001000
		Ligohe Plant	2807	1	36	0	0	5	0	00001001	00001001
		Panjin Storage	2808	1	0	0	0	5	5	00001000	00001000
		Yinkou Paper Plant	2809	1	0	0	0	5	8	00001000	00001000
		Nanheyan Irr. Sta.	2810	1	0	0	0	5	5	00001001	00001001
		Shuiyuan Commune	2811	1	0	0	0	5	5	00001001	00001001
		Yinkou Gate	2812	1	0	0	0	5	5	00001001	00001001
		Panjin Ch. Fertil.	2813	1	0	0	0	8	8	00000001	00000001
		Yinkou Glass Fiber	2814	1	0	0	0	8	8	00000001	00000001
		Shuang Tai Zi R.	2815	0	34	0	34	8	8	00000001	00000001
Guatemala	1976 2/4	Amatitlan 1	2901	1	0	0	0	7	9	00000011	00000011
		Amatitlan 2	2902	0	0	38	38	7	9	00000011	00000011
		Amatitlan 3	2903	0	37	0	37	0	9	00000000	00000000
		Amatitlan 4	2904	0	37	0	37	7	9	00000010	00000000
		Amatitlan 2	2905	0	0	38	38	7	0	00000001	00000001
		Amatitlan 3,4	2906	0	37	0	37	8	0	00000001	00000001
Tangshan	1976 7/28	Weigezhuang	3001	1	0	0	0	5	5	00001001	00001001
		Lujiato Mine	3002	1	0	0	0	5	5	00001001	00001001
		Lutai 3	3003	1	0	0	0	7	0	00000010	00000010
		Lutai 51	3004	1	0	0	0	7	0	00000010	00000010
		Lutai 52	3005	0	0	0	0	7	0	00000010	00000010
		Tangshan City	3006	0	0	0	0	8	8	00000001	00000001
		Qing Yin	3007	1	0	0	0	8	8	00000001	00000001
		Luan Nan	3008	1	0	0	0	8	8	00000001	00000001
		Le Ting	3009	1	0	0	0	8	8	00000001	00000001
		Coastal Region	3010	1	0	0	0	8	9	00000001	00000001
		Yao Yuan Village	3011	1	0	0	0	8	9	00000001	00000001
		Ma Feng	3012	0	0	0	0	8	9	00000001	00000001
		Wang Zhuang	3013	1	0	0	0	8	9	00000001	00000001
Vrancea	1977 3/4	Dimbovitza-1	3101	1	38	0	0	0	9	00000000	00000000
		Dimbovitza-2	3102	0	38	0	0	0	9	00000000	00000000
San Juan Argentina	1977 11/3	Barrio Castro B-1	3201	1	39	0	0	0	9	00000001	00000001
		Barrio Castro B-2	3202	1	39	0	0	0	9	00000000	00000000
		Caucete B-3	3203	1	0	0	0	0	9	00000001	00000001
		West of River B-4	3204	0	0	0	0	0	9	00000001	00000001
		West of River B-5	3205	0	0	0	0	0	9	00000001	00000001
		Fin. Santiago B-6	3206	1	0	0	0	0	9	00000001	00000001
		Escuela Normal 1	3207	1	55	0	0	0	9	00000000	00000000
		Escuela Normal 2	3208	1	55	0	0	0	9	00000000	00000000
		Escuela Normal 3	3209	1	55	0	0	0	9	00000000	00000000
		Escuela Normal 4	3210	1	55	0	0	0	9	00000000	00000000
		Escuela Normal 5	3211	1	55	0	0	0	9	00000000	00000000
		Escuela Normal 6	3212	1	55	0	0	0	9	00000000	00000000
		Airport (Rt 20) 1	3213	0	56	0	0	0	9	00000000	00000000
		Airport (Rt 20) 2	3214	0	56	0	0	0	9	00000000	00000000

Table A.3 (Continued)

EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	GCAS	BCAT	ECAT	REP	TAB
		Santa Rosa 1	3215	0	57	0	0	0	9	00000000	00000000
		Santa Rosa 2	3216	0	57	0	0	0	9	00000000	00000000
		Santa Rosa 3	3217	0	57	0	0	0	9	00000000	00000000
Izu-Oshima-Kinkai	1978 1/14	Mochikoshi	3301	1	0	0	0	7	7	00000011	00000010
Miyagiken-Oki - 1	1978 2/20	Arahama	3401	0	49	0	0	7	8	00000011	00000011
		Nakamura 1	3402	0	40	0	0	0	8	00000000	00000000
		Nakamura 4	3403	1	40	0	0	7	8	00000011	00000011
		Nakamura 5	3404	0	40	0	0	7	8	00000011	00000011
		Yuriageka 1	3405	0	41	0	0	0	8	00000011	00000011
		Yuriageka 2	3406	0	41	0	0	7	8	00000011	00000011
		Yuriageka 3	3407	0	41	0	0	0	8	00000000	00000000
		Yuriage Bridge 1	3408	0	42	0	0	7	8	00000011	00000011
		Yuriage Bridge 2	3409	0	42	0	0	7	8	00000011	00000011
		Yuriage Bridge 3	3410	0	42	0	0	7	8	00000011	00000011
		Yuriage Bridge 5	3411	0	42	0	0	0	8	00000000	00000000
		Oiira 1	3412	0	43	0	0	7	8	00000011	00000011
		Oiira 2	3413	0	43	0	0	7	8	00000011	00000011
		Kitawabu 2	3414	0	44	0	0	7	8	00000011	00000011
		Kitawabu 3	3415	0	44	0	0	0	8	00000000	00000000
		Shiomi 2	3416	0	45	0	0	0	8	00000000	00000000
		Shiomi 6	3417	0	45	0	0	7	8	00000011	00000011
		Hiyori 5	3418	0	46	0	0	0	8	00000000	00000000
		Hiyori 18	3419	0	46	0	0	7	8	00000011	00000011
		Nakajima 2	3420	0	47	0	0	0	8	00000000	00000000
		Nakajima 18	3421	0	47	0	0	7	8	00000011	00000011
		Sendaikou 1	3422	0	48	0	0	0	8	00000000	00000000
		Sendaikou 4	3423	0	48	0	0	7	8	00000011	00000011
		Ishinomaki 2	3424	0	50	50	0	7	8	00000011	00000011
		Ishinomaki 4	3425	0	50	51	0	0	8	00000000	00000000
Miyagiken-Oki - 2	1978 6/12	Arahama	3501	1	49	0	0	7	8	00000011	00000011
		Nakamura 1	3502	0	40	0	0	7	8	00000011	00000011
		Nakamura 4	3503	1	40	0	0	7	8	00000011	00000011
		Nakamura 5	3504	1	40	0	0	7	8	00000011	00000011
		Yuriageka 1	3505	1	41	0	0	7	8	00000011	00000011
		Yuriageka 2	3506	1	41	0	0	7	8	00000011	00000011
		Yuriageka 3	3507	0	41	0	0	7	8	00000000	00000000
		Yuriage Bridge 1	3508	1	42	0	0	7	8	00000011	00000011
		Yuriage Bridge 2	3509	1	42	0	0	7	8	00000011	00000011
		Yuriage Bridge 3	3510	1	42	0	0	7	8	00000011	00000011
		Yuriage Bridge 5	3511	0	42	0	0	7	8	00000011	00000011
		Oiira 1	3512	1	43	0	0	7	8	00000011	00000011
		Oiira 2	3513	1	43	0	0	7	8	00000011	00000011
		Kitawabu 2	3514	1	44	0	0	7	8	00000011	00000011
		Kitawabu 3	3515	0	44	0	0	7	8	00000011	00000011
		Shiomi 2	3516	0	45	0	0	7	8	00000011	00000011
		Shiomi 6	3517	1	45	0	0	7	8	00000011	00000011
		Hiyori 5	3518	0	46	0	0	7	8	00000011	00000011
		Hiyori 18	3519	1	46	0	0	7	8	00000011	00000011
		Nakajima 2	3520	0	47	0	0	7	8	00000011	00000011
		Nakajima 18	3521	1	47	0	0	7	8	00000011	00000011
		Sendaikou 1	3522	0	48	0	0	7	8	00000011	00000011
		Sendaikou 4	3523	0	48	0	0	7	8	00000011	00000011
		Ishinomaki 2	3524	1	50	50	0	7	8	00000011	00000011
		Ishinomaki 4	3525	0	50	51	0	7	8	00000011	00000011
Thessaloniki	1978 6/20	Greek Church	3601	0	0	0	0	0	9	00000000	00000000
		White Tower	3602	0	0	0	0	0	9	00000000	00000000
Guerrero	1979 3/14	Enmedio Zone 1	3701	0	0	0	0	0	9	00000000	00000001
		Enmedio Zone 2	3702	1	0	0	0	0	9	00000000	00000000
Montenegro	1979 4/15	Boca Kotorska	3801	1	0	0	0	0	9	00000000	00000000

Table A.3 (Continued)

EARTHQUAKE	DATE	SITE NAME	CASE	LIQ1	SITE	BORC	GCAS	BCAT	ECAT	REP	TAB
Imperial Valley	1979 10/15	Heber Rd. 1/A1	3901	0	52	79	0	7	9	00000011	00000011
		Heber Rd. 4,6/A2	3902	1	52	80	0	7	9	00000011	00000011
		Heber Rd. 7/A3	3903	0	52	81	0	7	9	00000011	00000011
		River Park A	3904	1	53	82	0	7	9	00000011	00000011
		River Park C	3905	1	53	83	0	7	9	00000011	00000011
		Wildlife 1Ns	3906	0	58	58	58	0	9	00000000	00000000
		Wildlife 2Ng1	3907	0	58	59	58	0	9	00000000	00000000
		Wildlife 2Ng3	3908	0	58	60	58	0	9	00000000	00000000
		Wildlife 3Ns	3909	0	58	61	58	0	9	00000000	00000000
		Wildlife 3Ns	3910	0	58	62	58	0	9	00000000	00000000
		Wildlife 5Ng	3911	0	58	63	58	0	9	00000000	00000000
		Vail V2a-b	3912	0	59	64	0	0	9	00000000	00000000
		Vail TV2a-b	3913	0	59	65	0	0	9	00000000	00000000
		Kornbloom K3	3914	0	60	66	60	0	9	00000000	00000000
		Kornbloom TK4a	3915	0	60	67	60	0	9	00000000	00000000
		Kornbloom TK4b	3916	0	60	68	60	0	9	00000000	00000000
		Kornbloom SK4-5	3917	0	60	69	60	0	9	00000000	00000000
		Radio Tower R2	3918	1	61	70	0	0	9	00000001	00000001
		Radio Tower R3	3919	1	61	71	0	0	9	00000001	00000001
		McKim TM6-7	3920	1	62	74	62	0	9	00000000	00000000
		McKim SM7	3921	1	62	75	62	0	9	00000000	00000000
		SNorthend SN2a	3922	0	63	76	0	0	9	00000000	00000000
		SNorthend SN2b	3923	0	63	77	0	0	9	00000000	00000000
		Young Y5	3924	0	0	78	0	0	9	00000000	00000000
		McKim Ranch A	3925	1	62	84	62	0	9	00000001	00000001
Mexicali Valley	1980 6/9	Town of Delta	4001	1	0	0	0	0	9	00000000	00000000
Mid-Chiba	1980 9/25	Owi Island 1	4101	0	54	54	54	7	7	00000011	00000011
		Owi Island 1	4102	0	54	54	54	8	0	00000011	00000011
		Owi Island 1	4103	0	54	54	55	7	7	00000011	00000011
		Owi Island 1	4104	0	54	54	55	8	0	00000011	00000011
Westmoreland	1981 4/26	Heber Rd. 1/A1	4201	0	52	79	0	7	9	00000010	00000010
		Heber Rd. 4,6/A2	4202	0	52	80	0	7	9	00000010	00000010
		Heber Rd. 7/A3	4203	0	52	81	0	7	9	00000010	00000010
		River Park A	4204	0	53	82	0	7	9	00000011	00000011
		River Park C	4205	0	53	83	0	7	9	00000011	00000011
		Wildlife 1Ns	4206	1	58	58	58	0	9	00000000	00000000
		Wildlife 2Ng1	4207	1	58	59	58	0	9	00000000	00000000
		Wildlife 2Ng3	4208	1	58	60	58	0	9	00000000	00000000
		Wildlife 3Ns	4209	1	58	61	58	0	9	00000000	00000000
		Wildlife 3Ns	4210	1	58	62	58	0	9	00000000	00000000
		Wildlife 5Ng	4211	0	58	63	58	0	9	00000000	00000000
		Vail V2a-b	4212	0	59	64	0	0	9	00000000	00000000
		Vail TV2a-b	4213	1	59	65	0	0	9	00000000	00000000
		Kornbloom K3	4214	1	60	66	60	0	9	00000000	00000000
		Kornbloom TK4a	4215	1	60	67	60	0	9	00000000	00000000
		Kornbloom TK4b	4216	1	60	68	60	0	9	00000000	00000000
		Kornbloom SK4-5	4217	1	60	69	60	0	9	00000000	00000000
		Radio Tower R2	4218	1	61	70	0	0	9	00000000	00000001
		Radio Tower R3	4219	1	61	71	0	0	9	00000000	00000001
		McKim TM6-7	4220	0	62	74	62	0	9	00000000	00000000
		McKim SM7	4221	0	62	75	62	0	9	00000000	00000000
		SNorthend SN2a	4222	0	63	76	0	0	9	00000000	00000000
		SNorthend SN2b	4223	0	63	77	0	0	9	00000000	00000000
		Young Y5	4224	0	0	78	0	0	9	00000000	00000000
		Wildlife B	4225	0	58	0	58	8	0	00000001	00000001
		Kornbloom B	4226	0	60	0	60	8	0	00000001	00000001
		McKim Ranch A	4227	0	62	84	62	8	0	00000001	00000001

Table A.4

Data From Previously Published Catalogs

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
0101	2	0.	-1.	6.6	-1.	-1.	-1.	39.	20.	.12	.14	-1.	.9	6.1	-1.	.55	6.	8.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0102	2	0.	-1.	6.6	-1.	-1.	-1.	39.	20.	.12	.14	-1.	.9	6.1	1.07	.55	12.	16.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0201	2	0.	-1.	6.1	-1.	-1.	-1.	47.	12.	.08	.09	-1.	.9	6.1	-1.	.55	6.	8.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0202	2	0.	-1.	6.1	-1.	-1.	-1.	47.	12.	.08	.09	-1.	.9	6.1	-1.	.55	12.	16.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0301	2	1.	-1.	8.4	-1.	-1.	-1.	32.	75.	.35	.39	-1.	.9	13.7	-1.	1.30	17.	15.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0301	4	1.	-1.	-2.	-1.	30.	-1.	-1.	-1.	-1.	-2.	-1.	.8	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
0301	7	1.	1.0	7.9	-1	-1.	-1.	-1.	-1.	.32	-1.	.364	.8	10.0	1.90	.98	20.	20.2	-1.	-1.	0.	-1.	5.	.46	2.3
0301	8	1.	-1.	7.9	-1	-1.	-1.	-1.	-1.	.32	.355	.375	.8	10.0	1.93	1.02	20.	19.5	1.30	25.5	0.	-1.	-1.	.46	-1.
0302	2	1.	-1.	8.4	-1.	-1.	-1.	32.	75.	.35	.37	-1.	1.8	9.1	-1.	1.00	10.	10.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0302	4	1.	-1.	-2.	-1.	30.	-1.	-1.	-1.	-1.	-1.	-1.	2.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
0302	7	1.	1.0	7.9	-1.	-1.	-1.	-1.	-1.	.32	-1.	.317	2.0	7.0	1.33	.83	10.	11.1	-1.	-1.	5.	-1.	0.	.28	2.2
0302	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.32	.31	.33	2.0	7.0	1.35	.86	10.	10.5	1.17	12.5	5.	-1.	-1.	.28	-1.
0303	2	0.	-1.	8.4	-1.	-1.	-1.	32.	75.	.35	.35	-1.	1.8	7.6	-1.	.85	19.	21.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0303	4	0.	-1.	-2.	-1.	30.	-1.	-1.	-1.	-1.	-1.	-1.	1.9	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
0303	7	1.	1.0	7.9	-1.	-1.	-1.	-1.	-1.	.28	-1.	.275	1.9	6.0	1.14	.73	17.	20.2	-1.	-1.	3.	-1.	26.	.8	3.0
0303	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.32	.315	.33	1.9	6.0	1.17	.75	17.	19.0	1.30	25.0	3	-1.	-1.	.8	-1.
0304	2	1.	-1.	8.4	-1.	-1.	-1.	32.	75.	.35	.35	-1.	2.4	6.1	-1.	.75	16.	19.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0304	4	1.	-1.	-2.	-1.	30.	-1.	-1.	-1.	-1.	-1.	-1.	2.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
0304	7	1.	1.0	7.9	-1.	-1.	-1.	-1.	-1.	.28	-1.	.246	2.4	5.0	.95	.69	13.	15.9	-1.	-1.	4.	-1.	24.	.60	3.3
0304	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.32	.275	.29	2.4	5.0	1.00	.72	13.	15.0	1.17	17.0	4.	-1.	-1.	.60	-1.
0305	3	0.	-1.	8.4	-1.	-1.	-1.	-1.	-1.	.35	-1.	-1.	1.8	12.2	-1.	-1.	-1.	48.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0306	6	1.	-1.	8.4	-1.	55.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.74	-1.	15.4	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0307	6	1.	-1.	8.4	-1.	51.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.45	-1.	20.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0401	6	1.	-1.	7.5	-1.	79.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.62	-1.	10.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0402	6	1.	-1.	7.5	-1.	65.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.37	-1.	10.7	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0403	6	0.	-1.	7.5	-1.	58.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.54	-1.	12.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0404	6	1.	-1.	7.5	-1.	37.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.39	-1.	4.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0405	6	1.	-1.	7.5	-1.	13.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.47	-1.	7.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0501	3	1.	-1.	8.3	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	2.4	4.6	-1.	.60	16.	20.4	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0501	4	1.	-1.	-2.	25.	-1.	15.	-1.	-1.	-1.	-1.	-1.	2.4	4.6	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0502	3	1.	-1.	8.3	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	2.4	7.6	-1.	.85	16.	17.4	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0502	4	1.	-1.	-2.	25.	-1.	15.	-1.	-1.	-1.	-1.	-1.	2.4	7.6	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0503	3	1.	-1.	8.3	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	2.4	4.9	-1.	.98	52.	52.2	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0504	3	1.	-1.	8.3	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	1.5	4.6	-1.	.52	24.	33.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0504	4	1.	-1.	-2.	-1.	-1.	15.	-1.	-1.	-1.	-1.	-1.	1.5	4.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
0505	6	1.	-1.	8.3	-1.	48.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.48	-1.	8.7	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0505	6	1.	-1.	8.3	-1.	48.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.69	-1.	5.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0506	6	1.	-1.	8.3	-1.	184.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.43	-1.	10.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0601	6	0.	-1.	6.9	-1.	51.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.74	-1.	15.4	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0602	6	0.	-1.	6.9	-1.	57.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.45	-1.	20.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0701	7	1.	0.7	7.9	-1.	-1.	-1.	-1.	-1.	.20	.163	-1.	4.0	7.0	1.25	.95	10.	10.3	-1.	-1.	10.	-1.	0.	.25	3.4
0701	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.20	.16	.17	4.0	7.0	1.28	.98	10.	10.0	1.09	11.0	10.	-1.	-1.	.25	-1.
0702	7	1.	0.7	7.9	-1.	-1.	-1.	-1.	-1.	.20	-1.	.168	4.0	8.0	1.44	1.04	1.	1.0	-1.	-1.	14.	5.	0.	.25	-1.
0703	7	1.	0.7	7.9	-1.	-1.	-1.	-1.	-1.	.20	-1.	.130	4.0	4.3	.73	.71	2.2	2.7	-1.	-1.	22.	11.	3.	.18	45.
0703	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.20	.13	.14	4.0	4.3	.76	.73	2.	2.5	1.09	2.5	22.	-2.	-2.	.18	-1.
0704	7	1.	1.0	7.9	-1.	-1.	-1.	-1.	-1.	.20	-1.	.225	1.0	8.0	1.52	.82	16.5	18.5	-1.	-1.	1.	-1.	0.	.28	1.6
0704	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.20	.225	.24	1.0	8.0	1.58	.85	16.	17.0	1.21	20.5	1.	-1.	-1.	.28	-1.
0705	7	1.	0.7	7.9	-1.	-1.	-1.	-1.	-1.	.20	-1.	.187	1.0	5.0	.95	.65	11.9	15.0	-1.	-1.	5.	-1.	0.	.3	2.7
0705	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.20	.22	.23	1.0	5.0	1.00	.67	12.	15.5	1.09	16.5	5.	-1.	-1.	.30	-1.
0706	7	0.	0.5	7.9	-1.	-1.	-1.	-1.	-1.	.20	-1.	.162	3.0	5.0	.95	.75	5.7	6.7	-1.	-1.	20.	5.	0.	.2	13.0
0706	8	0.	.5	7.9	-1.	-1.	-1.	-1.	-1.	.20	.16	.17	3.0	5.0	1.00	.78	6.	6.5	1.09	7.0	20.	-1.	-1.	.20	-1.
0707	7	0.	0.5	7.9	-1.	-1.	-1.	-1.	-1.	.20	-1.	.181	3.0	8.0	1.52	1.02	2.	2.0	-1.	-1.	33.	9.	0.	.15	25.0
0708	6	0.	-1.	7.9	-1.	114.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.62	-1.	10.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0709	6	0.	-1.	7.9	-1.	102.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.37	-1.	10.7	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0710	6	0.	-1.	7.9	-1.	105.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.54	-1.	12.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0711	6	1.	-1.	7.9	-1.	96.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.39	-1.	4.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0712	6	1.	-1.	7.9	-1.	77.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.47	-1.	7.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0713	6	1.	-1.	7.9	-1.	72.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.42	-1.	6.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0714	6	1.	-1.	7.9	-1.	75.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.47	-1.	15.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0801	2	1.	-1.	6.3	-1.	-1.	-1.	11.	15.	.20	.16	-1.	4.5	7.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0801	5	1.	-1.	6.3	-1.	11.	-1.	-1.	15.	.20	-1.	-1.	4.5	7.5	-1.	-1.	3.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0801	6	1.	1.0	6.3	-1.	11.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	1.06	-1.	2.9	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0901	6	1.	-1.	7.0	-1.	20.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.62	-1.	10.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0902	6	1.	-1.	7.0	-1.	19.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.37	-1.	10.7	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0903	6	0.	-1.	7.0	-1.	34.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.54	-1.	12.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0904	6	1.	-1.	7.0	-1.	51.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.39	-1.	4.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0905	6	0.	-1.	7.0	-1.	65.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.47	-1.	7.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
1001	3	0.	-1.	6.3	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	3.0	-2.	-1.	-2.	-2.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1002	3	0.	-1.	6.3	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	5.5	-2.	-1.	-2.	-2.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1003	3	0.	-1.	6.3	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	3.0	-2.	-1.	-2.	-2.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1004	3	0.	-1.	6.3	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	3.0	-2.	-1.	-2.	-2.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1005	7	0.	-1.	6.3	-1.	-1.	-1.	-1.	-1.	.20	-1.	.163	-1.	-1.	-1.	-1.	-1.	7.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1005	8	0.	-1.	6.3	-1.	-1.	-1.	-1.	-1.	.20	.205	.165	1.8	8.2	1.61	.90	8.	8.0	1.00	8.0	25.	-1.	-1.	.11	-1.
1006	8	0.	-1.	6.3	-1.	-1.	-1.	-1.	-1.	.16	.155	.125	1.8	6.1	1.11	.71	7.	8.5	1.00	8.5	2.	-1.	-1.	.44	-1.
1101	2	1.	-1.	7.0	-1.	-1.	-1.	8.	30.	.25	.155	-1.	4.6	4.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1101	5	1.	-1.	-2.	-1.	-1.	8.	-1.	-1.	.25	-1.	-1.	4.5	4.5	-1.	-1.	9.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1101	6	1.	-1.	7.0	-1.	8.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.77	-1.	9.8	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1102	2	1.	-1.	7.0	-1.	-1.	-1.	8.	30.	.25	.155	-1.	6.1	7.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1102	5	1.	-1.	7.0	-1.	-1.	8.	-1.	-1.	.25	-1.	-1.	6.0	7.5	-1.	-1.	4.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1102	6	1.	-1.	7.0	-1.	8.	-1.	-1.	-1.	.25	-1.	-1.	-1.	-1.	-1.	1.17	-1.	3.8	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1103	2	1.	-1.	7.0	-1.	-1.	-1.	8.	30.	.25	.26	-1.	1.5	6.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1103	5	1.	-1.	7.0	-1.	-1.	8.	-1.	-1.	.25	-1.	-1.	1.5	6.0	-1.	-1.	1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1103	6	1.	-1.	7.0	-1.	8.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.69	-1.	1.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1201	2	1.	-1.	8.3	-1.	-1.	-1.	161.	70.	.08	.08	-1.	1.5	4.0	-1.	.50	4.	6.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1201	4	1.	-1.	-2.	-1.	165.	-1.	-1.	-1.	-1.	-1.	-1.	2.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
1201	7	1.	1.0	8.0	-1.	-1.	-1.	-1.	-1.	.20	-1.	.193	2.0	5.0	.91	.61	8.	10.4	-1.	-1.	10.	-1.	14.	.4	5.0
1201	8	1.	-1.	8.0	-1.	-1.	-1.	-1.	-1.	.20	.18	.195	2.1	5.2	1.00	.69	8.	9.5	1.17	11.0	10.	-1.	-1.	.40	-1.
1202	2	1.	-1.	8.3	-1.	-1.	-1.	161.	70.	.08	.09	-1.	.6	2.4	-1.	.25	1.	2.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1202	4	1.	-1.	-2.	-1.	165.	-1.	-1.	-1.	-1.	-1.	-1.	.5	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
1202	7	1.	1.0	8.0	-1.	-1.	-1.	-1.	-1.	.20	-1.	.243	.5	3.5	.66	.36	1.	1.6	-1.	-1.	-1.	-1.	-1.	.20	-1.
1202	8	1.	-1.	8.0	-1.	-1.	-1.	-1.	-1.	.20	.225	.24	.6	3.7	.70	.40	1.	1.5	1.17	1.5	27.	-1.	-1.	20.	-1.
1203	4	1.	-1.	-2.	-1.	165.	-1.	-1.	-1.	-1.	-1.	-1.	2.5	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
1203	7	1.	1.0	8.0	-1.	-1.	-1.	-1.	-1.	.20	-1.	.148	2.5	3.0	.51	.46	2.	2.9	-1.	-1.	30.	4.	0.	.15	7.6
1203	8	1.	-1.	8.0	-1.	-1.	-1.	-1.	-1.	.20	.145	.155	2.4	3.0	.55	.49	2.	2.5	1.17	3.0	30.	-1.	-1.	.15	-1.
1204	7	1.	1.0	8.0	-1.	-1.	-1.	-1.	-1.	.16	-1.	.161	2.0	7.0	1.33	.83	10.	11.1	-1.	-1.	5.	-1.	0.	.28	2.2
1204	8	1.	-1.	8.0	-1.	-1.	-1.	-1.	-1.	.16	.155	.165	2.1	7.0	1.35	.86	10.	10.5	1.17	12.5	5.	-1.	-1.	.28	-1.
1301	7	0.	1.0	7.3	-1.	-1.	-1.	-1.	-1.	.35	-1.	.392	.8	4.0	.68	.38	7.	11.2	-1.	-1.	35.	14.	0.	.13	50.
1302	2	0.	-1.	7.2	-1.	-1.	-1.	6.	30.	.30	.32	-1.	.9	7.0	-1.	.70	28.	34.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1302	4	0.	-1.	-2.	-1.	5.	-1.	-1.	-1.	-1.	-1.	-1.	.8	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
1302	7	0.	0.0	7.3	-1.	-1.	-1.	-1.	-1.	.35	-1.	.388	.8	8.0	1.44	.72	29.	34.7	-1.	-1.	2.	-1.	0.	.8	2.4
1302	8	0.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.35	.395	.385	.9	8.2	1.57	.83	29.	31.5	1.30	40.5	2.	-1.	-1.	.80	-1.
1303	2	1.	-1.	7.2	-1.	-1.	-1.	6.	30.	.30	.30	-1.	3.4	7.0	-1.	.95	18.	18.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1303	4	1.	-1.	-2.	-1.	5.	-1.	-1.	-1.	-1.	-1.	-1.	3.7	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
1303	7	1.	1.0	7.3	-1.	-1.	-1.	-1.	-1.	.35	-1.	.267	3.7	7.0	1.26	.93	19.	19.8	-1.	-1.	4.	-1.	5.	.65	3.8
1303	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.35	.29	.28	3.7	7.0	1.35	1.01	19.	18.5	1.30	24.0	4.	-1.	-1.	.65	-1.
1304	2	1.	-1.	7.2	-1.	-1.	-1.	6.	30.	.30	.29	-1.	1.2	3.0	-1.	.38	3.	5.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1304	4	1.	-1.	-2.	-1.	5.	-1.	-1.	-1.	-1.	-1.	-1.	1.2	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
1304	7	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.40	-1.	.375	1.2	4.0	.76	.48	8.	11.5	-1.	-1.	0.	-1.	0.	.45	1.4
1304	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.40	.395	.38	1.2	4.0	.76	.49	8.	11.0	1.17	13.0	0.	-1.	-1.	.45	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
1305	2	1.	-1.	7.2	-1.	-1.	-1.	6.	30.	.30	.33	-1.	.9	6.1	-1.	.60	5.	7.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1305	4	1.	-1.	-2.	-1.	5.	-1.	-1.	-1.	-1.	-1.	-1.	.9	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
1305	7	0.	0.5	7.3	-1.	-1.	-1.	-1.	-1.	.40	-1.	.450	.9	6.0	1.02	.52	8.	11.1	-1.	-1.	21.	-1.	0.	.10	1.7
1306	7	0.	0.5	7.3	-1.	-1.	-1.	-1.	-1.	.40	-1.	.451	.9	7.5	1.31	.65	20.	25.2	-1.	-1.	0.	-1.	3.	.45	2.1
1306	8	0.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.40	.46	.45	.9	7.6	1.42	.75	20.	22.5	1.30	29.0	0.	-1.	-1.	.45	-1.
1401	3	0.	-1.	5.4	-1.	-1.	40.	-1.	-1.	-1.	-1.	-1.	2.4	17.1	-1.	1.50	22.	17.9	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1501	3	0.	-1.	5.5	-1.	-1.	11.	-1.	-1.	-1.	-1.	-1.	4.6	6.1	-1.	.94	4.	4.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1502	2	1.	-1.	5.5	-1.	-1.	-1.	6.	18.	.18	.13	-1.	2.4	3.0	-1.	.50	7.	10.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1502	7	1.	1.0	5.5	-1.	-1.	-1.	-1.	-1.	.19	.10	-1.	-1.	-1.	.60	.49	6.	6.1	-1.	-1.	3.	-1.	-1.	.2	-1.
1502	8	1.	-1.	5.3	-1.	-1.	-1.	-1.	-1.	.19	.13	.085	2.4	3.0	.52	.46	5.	7.0	.75	5.5	3.	-1.	-1.	.34	-1.
1503	3	0.	-1.	5.5	-1.	-1.	10.	-1.	-1.	-1.	-1.	-1.	3.7	4.0	-1.	.68	14.	16.9	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1504	3	0.	-1.	5.5	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	2.4	4.6	-1.	.60	16.	20.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1505	3	0.	-1.	5.5	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	2.4	7.6	-1.	.85	16.	17.4	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1506	3	0.	-1.	5.5	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	1.5	6.1	-1.	.98	52.	52.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1507	3	0.	-1.	5.5	-1.	-1.	13.	-1.	-1.	-1.	-1.	-1.	1.5	4.6	-1.	.52	24.	33.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1508	3	0.	-1.	5.5	-1.	-1.	11.	-1.	-1.	-1.	-1.	-1.	1.5	6.1	-1.	.63	6.	7.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1509	3	0.	-1.	5.5	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	4.6	6.1	-1.	.94	20.	20.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1510	3	0.	-1.	5.5	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	2.4	4.6	-1.	.60	20.	25.7	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1511	3	0.	-1.	5.5	-1.	-1.	11.	-1.	-1.	-1.	-1.	-1.	.9	1.2	-1.	.19	4.	9.2	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1512	3	0.	-1.	5.5	-1.	-1.	11.	-1.	-1.	-1.	-1.	-1.	.9	1.2	-1.	.19	6.	13.9	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1513	3	0.	-1.	5.5	-1.	-1.	11.	-1.	-1.	-1.	-1.	-1.	1.2	3.7	-1.	.41	6.	9.4	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1514	3	0.	-1.	5.5	-1.	-1.	11.	-1.	-1.	-1.	-1.	-1.	1.2	4.3	-1.	.46	6.	8.8	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1515	3	0.	-1.	5.5	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	3.0	3.7	-1.	.60	11.	14.2	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1516	3	0.	-1.	5.5	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	3.0	4.0	-1.	.62	10.	12.7	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1517	3	0.	-1.	5.5	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	1.8	5.8	-1.	.64	12.	15.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1518	3	0.	-1.	5.5	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	1.2	5.8	-1.	.58	16.	20.9	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1519	3	0.	-1.	5.5	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	1.8	3.7	-1.	.48	10.	14.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1520	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	2.4	7.6	-1.	.85	3.	3.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1521	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	2.4	9.1	-1.	.97	5.	5.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1522	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	1.8	5.8	-1.	.64	7.	8.7	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1523	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	1.8	5.8	-1.	.64	5.	6.2	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1524	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	1.8	3.0	-1.	.43	9.	13.8	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1525	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	1.8	4.6	-1.	.55	5.	6.8	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1526	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	1.8	4.6	-1.	.50	8.	11.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
1527	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	1.8	4.6	-1.	.50	5.	7.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1528	3	0.	-1.	5.5	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	1.8	4.6	-1.	.50	15.	21.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1529	3	0.	-1.	5.5	-1.	-1.	22.	-1.	-1.	-1.	-1.	-1.	1.8	3.0	-1.	.43	3.	4.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1530	3	0.	-1.	5.5	-1.	-1.	22.	-1.	-1.	-1.	-1.	-1.	2.1	2.7	-1.	.43	7.	10.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1531	3	0.	-1.	5.5	-1.	-1.	22.	-1.	-1.	-1.	-1.	-1.	.6	1.5	-1.	.18	13.	30.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1532	3	0.	-1.	5.5	-1.	-1.	26.	-1.	-1.	-1.	-1.	-1.	1.2	4.3	-1.	.46	7.	10.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1533	3	0.	-1.	5.5	-1.	-1.	19.	-1.	-1.	-1.	-1.	-1.	1.2	4.6	-1.	.41	5.	7.8	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1534	3	0.	-1.	5.5	-1.	-1.	19.	-1.	-1.	-1.	-1.	-1.	1.2	3.7	-1.	.35	12.	20.2	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1601	2	1.	-1.	8.4	-1.	-1.	-1.	113.	75.	.15	.15	-1.	3.7	4.6	-1.	.75	6.	7.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1602	2	1.	-1.	8.4	-1.	-1.	-1.	113.	75.	.15	.15	-1.	3.7	4.6	-1.	.75	8.	10.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1603	2	0.	-1.	8.4	-1.	-1.	-1.	113.	75.	.15	.15	-1.	3.7	6.1	-1.	.90	15.	19.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1604	1	0.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.15	.136	-1.	3.7	7.0	-1.	-1.	10.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1604	5	0.	-1.	8.4	-1.	-1.	-1.	113.	-1.	.15	-1.	-1.	3.5	7.0	-1.	-1.	10.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1605	1	0.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.15	.140	-1.	3.7	7.9	-1.	-1.	35.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1605	5	0.	-1.	8.4	-1.	-1.	-1.	113.	-1.	.15	-1.	-1.	3.5	8.0	-1.	-1.	35.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1701	2	1.	-1.	8.3	-1.	-1.	-1.	97.	180.	.15	.18	-1.	.0	6.1	-1.	.55	5.	7.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1701	4	1.	-1.	-2.	-2.	142.	-1.	-1.	-1.	-1.	-1.	-1.	.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	10.	-1.	-1.	-1.	-1.
1701	7	1.	1.0	8.3	-1.	-1.	-1.	-1.	-1.	.15	-1.	.210	.0	6.0	1.14	.54	5.	4.9	-1.	-1.	40.	-1.	-1.	.2	-1.
1702	2	1.	-1.	8.3	-1.	-1.	-1.	97.	180.	.15	.15	-1.	2.4	6.1	-1.	.75	5.	6.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1702	4	1.	-1.	-2.	-2.	142.	-1.	-1.	-1.	-1.	-1.	-1.	-2.	-2.	-1.	-1.	-2.	-1.	-1.	-1.	10.	-1.	-1.	-1.	-1.
1702	7	1.	1.0	8.3	-1.	-1.	-1.	-1.	-1.	.15	-1.	.156	2.4	6.0	1.14	.73	5.	4.2	-1.	-1.	40.	-1.	-1.	-1.	-1.
1703	2	0.	-1.	8.3	-1.	-1.	-1.	113.	180.	.12	.146	-1.	.0	7.6	-1.	.75	-2.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1703	4	0.	-1.	-2.	-2.	145.	-1.	-1.	-1.	-1.	-1.	-1.	3.1	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1704	2	1.	-1.	8.3	-1.	-1.	-1.	89.	180.	.16	.185	-1.	.0	6.1	-1.	.55	10.	14.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1704	4	1.	-2.	-2.	-2.	126.	-1.	-1.	-1.	-1.	-1.	-1.	.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	10.	-1.	-1.	-1.	-1.
1704	7	1.	1.0	8.3	-1.	-1.	-1.	-1.	-1.	.16	-1.	.224	.0	6.0	1.14	.54	10.	9.8	-1.	-1.	10.	-1.	-1.	.2	-1.
1705	2	1.	-1.	8.3	-1.	-1.	-1.	56.	180.	.25	.25	-1.	1.5	6.1	-1.	.65	13.	16.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1705	4	1.	-1.	-2.	-1.	-1.	72.	1.	-1.	-1.	-1.	-1.	-2.	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1705	7	1.	1.0	8.3	-1.	-1.	-1.	1.	-1.	.25	-1.	.2[illegible]	1.5	6.0	1.10	.63	13.	11.9	-1.	-1.	0.	-1.	-1.	-1.	-1.
1706	1	1.	-1.	-1.	-1.	-1.	-1.	-1.	120.	.15	.192	-1.	.0	4.6	-1.	-1.	16.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1801	2	1.	-1.	7.5	-1.	-1.	-1.	52.	40.	.18	.195	-1.	.9	6.1	-1.	.60	6.	8.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1801	5	1.	-1.	7.5	-1.	-1.	-1.	52.	40.	.16	-1.	-1.	1.0	6.0	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1802	2	1.	-1.	7.5	-1.	-1.	-1.	52.	40.	.18	.195	-1.	.9	7.6	-1.	.75	15.	18.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1802	5	1.	-1.	7.5	-1.	-1.	-1.	52.	40.	.16	-1.	-1.	1.0	7.5	-1.	-1.	8.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1803	2	0.	-1.	7.5	-1.	-1.	-1.	52.	40.	.18	.195	-1.	.9	6.1	-1.	.60	12.	16.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1803	5	0.	-1.	7.5	-1.	-1.	-1.	52.	40.	.16	-1.	-1.	1.0	6.0	-1.	-1.	12.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1804	2	0.	-1.	7.5	-1.	-1.	-1.	52.	40.	.18	.12	-1.	3.7	7.6	-1.	1.00	6.	6.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CER	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	QC	D50	UC
1804	5	0.	-1.	7.5	-1.	-1.	-1.	52.	40.	.16	-1.	-1.	3.5	7.5	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1805	1	1.	-1.	-1.	-1.	-1.	-1.	-1.	20.	.16	.194	-1.	.9	6.1	-1.	-1.	7.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1805	5	1.	-1.	7.5	-1.	-1.	-1.	52.	20.	.16	-1.	-1.	1.0	6.0	-1.	-1.	7.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1806	1	0.	-1.	-1.	-1.	-1.	-1.	-1.	20.	.16	.156	-1.	3.7	8.8	-1.	-1.	7.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1806	5	0.	-1.	7.5	-1.	-1.	-1.	-1.	20.	.16	-1.	-1.	3.5	8.5	-1.	-1.	7.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1807	1	0.	-1.	-1.	-1.	-1.	-1.	-1.	20.	.16	.194	-1.	.9	6.1	-1.	-1.	12.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1807	5	0.	-1.	7.5	-1.	-1.	-1.	-1.	20.	.16	-1.	-1.	1.0	6.0	-1.	-1.	12.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1808	3	0.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	3.0	12.2	-1.	-1.	-1.	55.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1809	4	0.	-1.	7.5	-2.	51.	-1.	-1.	-1.	-1.	-1.	-1.	.9	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1810	4	1.	-1.	7.5	-2.	51.	-1.	-1.	-1.	-1.	-1.	-1.	.9	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1811	4	1.	-1.	7.5	-2.	51.	-1.	-1.	-1.	-1.	-1.	-1.	.9	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1812	4	1.	-1.	7.5	-2.	51.	-1.	-1.	-1.	-1.	-1.	-1.	.9	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1813	4	1.	-1.	7.5	-2.	51.	-1.	-1.	-1.	-1.	-1.	-1.	.9	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
1814	7	1.	1.0	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.170	1.0	7.0	1.33	.73	8.	9.5	-1.	-1.	2.	-1.	0.	.3	2.5
1814	8	1.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.16	.18	.18	.9	7.0	1.35	.74	8.	9.0	1.09	10.0	2.	-1.	-1.	.30	-1.
1815	7	0.	0.5	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.170	1.0	7.0	1.33	.73	12.	14.2	-1.	-1.	2.	-1.	0.	.3	2.5
1815	8	0.	0.5	7.5	-1.	-1.	-1.	-1.	-1.	.16	.18	.18	.9	7.0	1.35	.74	12.	13.5	1.09	15.0	2.	-1.	-1.	.30	-1.
1816	7	0.	0.0	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.151	2.0	7.0	1.31	.81	18.	20.3	-1.	-1.	2.	-1.	0.	.3	2.5
1816	8	0.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.18	.18	.18	1.8	7.0	1.35	.83	18.	19.5	1.21	23.5	2.	-1.	-1.	.30	-1.
1817	7	1.	1.0	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.168	1.0	10.0	1.90	1.00	10.	10.0	-1.	-1.	2.	-1.	0.	.3	2.5
1817	8	1.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.16	.18	.18	.9	10.1	1.93	1.02	10.	10.0	1.09	10.5	2.	-1.	-1.	.30	-1.
1818	7	0.	0.5	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.168	1.0	10.0	1.90	1.00	16.	16.0	-1.	-1.	2.	-1.	0.	.3	2.5
1818	8	0.	0.5	7.5	-1.	-1.	-1.	-1.	-1.	.16	.18	.18	.9	10.1	1.93	1.02	16.	15.5	1.09	17.0	2.	-1.	-1.	.30	-1.
1819	7	0.	0.0	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.134	2.0	10.0	1.88	1.08	20.	19.1	-1.	-1.	2.	-1.	0.	.3	2.5
1819	8	0.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.18	.185	.185	1.8	10.1	1.93	1.11	20.	19.0	1.21	22.5	2.	-1.	-1.	.30	-1.
1820	7	1.	1.0	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.205	.0	4.3	.82	.39	4.	6.2	-1.	-1.	10.	4.	0.	.4	6.0
1820	8	1.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.16	.21	.21	.0	4.3	.82	.40	4.	6.0	1.09	6.5	10.	-1.	-1.	.40	-1.
1821	7	0.	0.0	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.161	1.3	6.0	1.14	.67	27.	33.5	-1.	-1.	0.	-1.	0.	.3	1.4
1821	8	0.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.18	.195	.195	1.2	6.1	1.17	.68	27.	32.0	1.21	38.5	0.	-1.	-1.	.30	-1.
1822	7	0.	0.0	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.128	2.5	6.0	.95	.70	12.	14.6	-1.	-1.	0.	-1.	0.	.38	1.5
1822	8	0.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.18	.165	.165	2.4	6.1	1.17	.81	12.	13.0	1.09	14.5	0.	-1.	-1.	.36	-1.
1823	7	1.	1.0	7.5	-1.	-1.	-1.	-1.	-1.	.16	-1.	.177	.6	4.5	.86	.47	6.	8.7	-1.	-1.	0.	-1.	0.	.40	1.6
1823	8	1.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.16	.185	.185	.6	4.6	.88	.48	6.	8.5	1.09	9.0	0.	-1.	-1.	.40	-1.
1901	8	0.	0.0	4.9	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	2.4	17.1	-1.	1.50	22.	17.9	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2001	2	1.	-1.	6.3	-1.	-1.	-1.	56.	15.	.13	.085	-1.	.9	.9	-1.	.165	3.	5.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2001	8	1.	-1.	6.3	-1.	-1.	-1.	-1.	-1.	.13	.085	.07	.9	.9	.16	.16	3.	6.0	.56	3.5	-1.	-1.	-1.	-1.	-1.
2101		1.	-1.	6.3	-2.	8.	-1.	-1.	-1.	-1.	-1.	-1.	1.8	-2.	-1.	-1.	7.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2102	4	1.	-1.	6.3	-2.	8.	-1.	-1.	-1.	-1.	-1.	-1.	1.4	-2.	-1.	-1.	7.	-1.	-1.	-1.	20.	-1.	-1.	-1.	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	R	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
2201	1	1.	-1.	-1.	-1.	-1.	-1.	-1.	20.	.19	.237	-1.	.9	5.8	-1.	-1.	13.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2201	2	1.	-1.	7.8	-1.	-1.	-1.	161.	45.	.18	.205	-1.	.9	4.6	-1.	.60	6.	9.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2201	4	1.	-1.	-2.	20.	-1.	-1.	160.	-1.	-1.	-1.	-1.	1.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2201	5	1.	1.0	7.8	-1.	283.	-1.	-1.	45.	.18	-1.	-1.	1.0	4.5	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2201	6	1.	-1.	7.8	-1.	160.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.26	-1.	5.8	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2201	7	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.20	-1.	.214	1.0	4.0	.76	.46	5.	7.3	-1.	-1.	20.	6.	0.	.12	6.0
2201	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.20	.21	.22	.9	4.0	.76	.46	5.	7.0	1.17	8.0	20.	-1.	-1.	.12	-1.
2202	1	0.	-1.	-1.	-1.	-1.	-1.	-1.	20.	.21	.205	-1.	1.5	3.7	-1.	-1.	18.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2202	2	0.	-1.	7.8	-1.	-1.	-1.	-2.	45.	.21	.185	-1.	1.5	3.0	-1.	.40	15.	23.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2202	5	0.	0.0	7.8	-1.	172.	-1.	-1.	45.	.21	-1.	-1.	1.5	3.0	-1.	-1.	15.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2203	1	1.	-1.	-1.	-1.	-1.	-1.	-1.	20.	.21	.210	-1.	1.5	3.7	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2203	5	1.	-1.	7.8	-1.	172.	-1.	-1.	20.	.21	-1.	-1.	1.5	3.5	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2204	2	0.	-1.	7.8	-1.	-1.	-1.	-2.	45.	.21	.23	-1.	.9	3.7	-1.	.40	14.	21.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2204	5	0.	-1.	7.8	-1.	172.	-1.	-1.	45.	.21	-1.	-1.	1.0	3.5	-1.	-1.	14.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2205	2	1.	-1.	7.8	-1.	-1.	-1.	-2.	45.	.21	.185	-1.	1.5	2.1	-1.	.40	6.	9.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2205	5	1.	-1.	7.8	-1.	172.	-1.	-1.	45.	.21	-1.	-1.	2.0	3.5	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2206	4	0.	-1.	7.8	-2.	160.	-1.	-1.	-1.	-1.	-1.	-1.	1.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2207	4	0.	-1.	7.8	-2.	160.	-1.	-1.	-1.	-1.	-1.	-1.	1.5	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2207	7	0.	0.0	7.9	-1.	-1.	-1.	-1.	-1.	.23	-1.	.222	2.0	6.0	1.14	.74	28.	33.1	-1.	-1.	5.	-1.	0.	.25	-1.
2207	8	0.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.23	.215	.23	2.1	6.1	1.17	.78	28.	31.0	1.21	37.5	5.	-1.	-1.	.25	-1.
2208	4	0.	-1.	7.8	-2.	160.	-1.	-1.	-1.	-1.	-1.	-1.	1.3	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2208	7	0.	0.0	7.9	-1.	-1.	-1.	-1.	-1.	.23	-1.	.246	1.0	4.0	.76	.46	16.	23.4	-1.	-1.	5.	-1.	0.	.25	-1.
2208	8	0.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.23	.24	.255	.9	4.0	.76	.46	16.	22.5	1.21	27.5	5.	-1.	-1.	.25	-1.
2209	4	0.	-1.	7.8	-2.	160.	-1.	-1.	-1.	-1.	-1.	-1.	1.5	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2210	4	0.	-1.	7.8	-2.	160.	-1.	-1.	-1.	-1.	-1.	-1.	1.5	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2211	4	0.	-1.	7.8	-2.	160.	-1.	-1.	-1.	-1.	-1.	-1.	1.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2212	4	0.	-1.	7.8	-2.	160.	-1.	-1.	-1.	-1.	-1.	-1.	1.3	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-2.	-1.	-1.	-1.	-1.
2213	7	1.	1.0	7.9	-1.	-1.	-1.	-1.	-1.	.23	-1.	.270	.6	4.0	.76	.42	6.	9.1	-1.	-1.	5.	-1.	0.	.25	-1.
2213	8	1.	-1.	7.9	-1.	-1.	-1.	-1.	-1.	.23	.26	.275	.6	4.0	.76	.42	6.	9.0	1.09	9.5	5.	-1.	-1.	.25	-1.
2301	4	0.	-2.	6.1	50.	-1.	69.	-1.	-1.	-1.	-1.	-1.	8.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2302	4	0.	-2.	6.1	50.	-1.	69.	-1.	-1.	-1.	-1.	-1.	8.0	10.0	-1.	-1.	47.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2303	4	0.	-2.	6.1	50.	-1.	69.	-1.	-1.	-1.	-1.	-1.	2.0	6.3	-1.	-1.	10.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2304	4	0.	-2.	6.1	50.	-1.	69.	-1.	-1.	-1.	-1.	-1.	2.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2305	4	0.	-2.	6.1	50.	-1.	69.	-1.	-1.	-1.	-1.	-1.	3.5	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2306	4	0.	-2.	6.1	50.	-1.	79.	-1.	-1.	-1.	-1.	-1.	3.0	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2501	4	0.	-1.	7.1	-1.	-1.	130.	-1.	-1.	-1.	-1.	-1.	3.7	7.	-1.	-1.	12.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2601	2	1.	-1.	6.6	-1.	-1.	-1.	8.	15.	.40	.28	-1.	4.6	6.1	-1.	1.00	2.	2.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2601	7	1.	1.0	6.6	-1.	-1.	-1.	-1.	-1.	.40	-1.	.237	4.5	6.0	1.14	.98	2.	1.4	-1.	-1.	10.	-1.	-1.	-1.	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
2601	8	1.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.45	.325	.28	4.6	6.1	1.14	.98	2.	2.0	.75	1.5	50.	-1.	-1.	-1.	-1.
2602	2	1.	-1.	6.6	-1.	-1.	-1.	8.	15.	.45	.16	-1.	16.8	16.8	-1.	3.30	24.	11.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2602	4	1.	-1.	6.6	-2.	-1.	24.	-1.	-1.	-1.	-1.	-1.	-2.	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2602	7	1.	1.0	6.6	-1.	-1.	-1.	-1.	-1.	.45	-1.	.190	16.5	16.5	3.22	3.22	24.	7.4	-1.	-1.	50.	-1.	-1.	.07	-1.
2603	7	1.	1.0	6.6	-1.	-1.	-1.	-1.	-1.	.45	-1.	.287	3.0	6.0	1.10	.88	9.	6.9	-1.	-1.	20.	-1.	-1.	.1	-1.
2603	8	1.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.45	.345	.30	3.0	6.1	1.10	.88	9.	9.5	.75	7.0	20.	-1.	-1.	.1	-1.
2604	3	0.	-1.	6.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	8.3	19.8	-1.	3.42	15.	8.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2605	3	0.	-1.	6.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	18.3	25.9	-1.	3.88	32.	16.2	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2606	3	0.	-1.	6.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	14.3	25.9	-1.	2.89	42.	24.7	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2607	3	0.	-1.	6.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	1.5	4.6	-1.	.48	8.	11.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2608	3	0.	-1.	6.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	16.8	19.8	-1.	3.43	21.	11.3	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2609	3	0.	-1.	6.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	10.7	15.2	-1.	2.28	30.	19.9	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2610	3	0.	-1.	6.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	10.7	16.8	-1.	2.61	26.	16.0	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2611	3	0.	-1.	6.6	-1.	-1.	24.	-1.	-1.	-1.	-1.	-1.	13.7	21.3	-1.	3.25	39.	21.6	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2701	4	0.	-1.	7.3	-1.	-1.	280.	-1.	-1.	-1.	-1.	-1.	3.1	-2.	-1.	-1.	-2.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2801	5	0.	-1.	7.3	-1.	90.	-1.	-1.	-1.	.075	-1.	-1.	1.0	8.0	-1.	-1.	15.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2802	5	0.	-1.	7.3	-1.	90.	-1.	-1.	-1.	.075	-1.	-1.	1.0	12.0	-1.	-1.	20.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2803	5	0.	-1.	7.3	-1.	90.	-1.	-1.	-1.	.075	-1.	-1.	2.0	8.0	-1.	-1.	9.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2803	8	0.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.10	.095	.095	2.0	8.0	1.14	.73	9.5	11.	1.00	11.0	-1.	-1.	-1.	-1.	-1.
2804	5	0.	-1.	7.3	-1.	90.	-1.	-1.	-1.	.075	-1.	-1.	2.0	12.0	-1.	-1.	15.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2805	5	0.	-1.	7.3	-1.	104.	-1.	-1.	-1.	.075	-1.	-1.	2.0	13.0	-1.	-1.	14.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2805	8	0.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.10	.095	.09	2.0	13.0	2.58	1.48	14.5	11.5	1.00	11.5	-1.	-1.	-1.	-1.	-1.
2806	5	1.	-1.	7.3	-1.	95.	-1.	-1.	-1.	.075	-1.	-1.	1.5	5.2	-1.	-1.	5.5	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2807	5	0.	-1.	7.3	-1.	95.	-1.	-1.	-1.	.075	-1.	-1.	1.5	6.2	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2807	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.10	.10	.10	1.5	6.2	1.22	.76	6.	6.5	1.00	6.5	-1.	-1.	-1.	-1.	-1.
2808	5	0.	-1.	7.3	-1.	57.	-1.	-1.	-1.	.075	-1.	-1.	1.5	7.0	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2809	5	0.	-1.	7.3	-1.	53.	-1.	-1.	-1.	.15	-1.	-1.	1.0	10.5	-1.	-1.	11.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2809	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.20	.21	.20	1.5	8.2	1.61	.94	11.	11.0	1.00	13.5	-1.	-1.	-1.	-1.	-1.
2810	5	1.	-1.	7.3	-1.	60.	-1.	-1.	-1.	.075	-1.	-1.	2.0	3.0	-1.	-1.	6.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2810	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.13	.10	.095	2.0	3.0	.58	.48	6.	8.5	.75	6.5	-1.	-1.	-1.	-1.	-1.
2811	5	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.15	-1.	-1.	2.0	10.0	-1.	-1.	9.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2811	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.20	.20	.195	2.0	10.0	1.98	1.18	9.	8.0	1.00	8.0	-1.	-1.	-1.	-1.	-1.
2812	5	1.	-1.	7.3	-1.	55.	-1.	-1.	-1.	.15	-1.	-1.	2.0	10.3	-1.	-1.	9.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
2812	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.20	.195	.19	2.0	10.3	2.04	1.21	9.	8.0	1.00	8.0	-1.	-1.	-1.	-1.	-1.
2813	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.13	.135	.13	1.5	9.1	1.79	1.03	8.	8.0	1.00	8.5	67.	-1.	-1.	.064	-1.
2814	8	1.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.20	.21	.20	1.5	8.2	1.61	.94	13.	13.5	1.00	13.5	48.	-1.	-1.	.078	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
2815	8	0.	-1.	7.3	-1.	-1.	-1.	-1.	-1.	.10	.105	.10	1.5	8.2	1.61	.94	9.	9.0	1.00	9.0	-1.	-1.	-1.	-1.	-1.
2901	7	1.	1.0	7.5	-1.	-1.	-1.	-1.	-1.	.135	-1.	.169	1.5	4.6	.60	.29	8.	9.3	-1.	-1.	3.	-1.	-1.	-1.	-1.
2901	8	1.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.135	.19	.19	1.5	10.4	1.42	.88	6.	6.0	.75	5.0	3.	-1.	-1.	.8	-1.
2902	7	0.	0.5	7.5	-1.	-1.	-1.	-1.	-1.	.135	-1.	.135	2.0	4.6	.58	.35	8.	9.3	-1.	-1.	3.	-1.	-1.	-1.	-1.
2902	8	0.	0.5	7.5	-1.	-1.	-1.	-1.	-1.	.135	.135	.135	2.4	4.6	.56	.34	8.	12.0	.75	9.0	3.	-1.	-1.	.8	-1.
2904	7	1.	1.0	7.5	-1.	-1.	-1.	-1.	-1.	.135	-1.	.128	3.3	7.0	1.03	.64	14.	12.7	-1.	-1.	3.	-1.	-1.	-1.	-1.
2905	8	1.	0.5	7.5	-1.	-1.	-1.	-1.	-1.	.135	.18	.18	2.4	11.6	1.57	.91	15.	15.5	.75	11.5	3.	-1.	-1.	.8	-1.
2906	8	0.	-1.	7.5	-1.	-1.	-1.	-1.	-1.	.135	.15	.15	-2.	10.7	1.40	.73	16.	18.5	.75	14.0	3.	-1.	-1.	.8	-1.
3001	5	1.	-1.	7.8	-1.	43.	-1.	-1.	-1.	.15	-1.	-1.	1.4	2.3	-1.	-1.	11.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
3001	8	1.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.20	.16	.17	1.4	2.3	.43	.33	11.	18.0	.75	13.5	-1.	-1.	-1.	-1.	-1.
3002	5	1.	-1.	7.8	-1.	25.	-1.	-1.	-1.	.30	-1.	-1.	1.0	7.0	-1.	-1.	4.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
3002	8	1.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.35	.39	.405	1.0	7.0	1.34	.74	4.	4.5	1.00	4.5	-1.	-1.	-1.	-1.	-1.
3003	7	1.	-1.	7.8	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
3004	7	1.	0.7	7.8	-1.	-1.	-1.	-1.	-1.	.20	-1.	.192	.5	11.0	2.20	1.30	-1.	7.0	-1.	-1.	50.	15.	-1.	.07	20.
3005	7	0.	0.5	7.8	-1.	-1.	-1.	-1.	-1.	.20	-1.	.192	.5	11.0	2.20	1.30	-1.	8.0	-1.	-1.	50.	15.	-1.	.07	20.
3006	8	0.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.50	.405	.42	3.0	5.3	1.00	.77	30.	33.5	1.00	33.5	10.	-1.	-1.	.196	-1.
3007	8	1.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.35	.38	.395	.9	5.3	1.04	.60	17.	21.5	1.00	21.5	20.	-1.	-1.	.137	-1.
3008	8	1.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.22	.165	.175	3.7	5.4	.99	.82	20.	22.0	1.00	22.0	3.	-1.	-1.	.17	-1.
3009	8	1.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.20	.155	.16	1.5	2.0	.37	.32	10.	16.5	.75	12.5	12.	-1.	-1.	.22	-1.
3010	8	1.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.13	.14	.145	1.2	6.1	1.19	.70	10.	11.5	1.00	11.5	12.	-1.	-1.	.14	-1.
3011	8	1.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.20	.22	.23	.9	6.1	1.20	.68	9.	11.0	1.00	11.0	-1.	-1.	-1.	-1.	-1.
3012	8	0.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.07	.06	.05	4.0	9.1	1.74	1.22	13.	11.5	1.00	11.5	1.	-1.	-1.	.30	-1.
3013	8	1.	-1.	7.6	-1.	-1.	-1.	-1.	-1.	.20	.205	.21	1.5	7.0	1.37	.82	11.	12.5	1.00	12.5	2.	-1.	-1.	.32	-1.
3201	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.20	.165	.16	4.6	8.2	1.45	1.08	9.	8.5	.75	6.0	20.	-1.	-1.	.15	-1.
3203	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.20	.165	.16	6.7	11.9	1.56	1.04	12.	11.5	.75	8.5	-1.	-1.	-1.	-1.	-1.
3204	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.20	.20	.195	1.2	3.7	.64	.40	14.	21.5	.75	16.0	4.	-1.	-1.	.29	-1.
3205	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.20	.155	.15	2.1	3.0	.54	.44	14.	19.5	.75	14.5	3.	-1.	-1.	.24	-1.
3206	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.20	.195	.19	1.8	5.2	.91	.58	6.	8.0	.75	6.0	51.	-1.	-1.	-1.	-1.
3301	7	1.	1.0	7.0	-1.	-1.	-1.	-1.	-1.	.25	-1.	.244	1.0	7.0	1.33	.73	-1.	1.2	-1.	-1.	90.	0.	0.	-1.	-1.
3401	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.10	-1.	.092	1.0	6.3	1.21	.68	10.	12.3	-1.	-1.	0.	-1.	0.	.60	-1.
3401	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.10	.11	.095	.9	6.4	1.23	.68	10.	12.0	1.09	13.0	0.	-1.	-1.	.60	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
3403	7	0.	0.7	6.7	-1.	-1.	-1.	-1.	-1.	.12	-1.	.117	.5	3.3	.63	.42	5.	8.1	-1.	-1.	5.	-1.	18.	.70	5.0
3403	8	1.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.12	.135	.115	.6	3.4	.64	.37	5.	8.0	1.00	8.0	5.	-1.	-1.	.70	-1.
3404	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.12	-1.	.098	1.3	3.3	.63	.42	7.	10.6	-1.	-1.	4.	-1.	2.	.28	2.5
3404	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.12	.115	.10	1.2	3.4	.64	.43	7.	10.0	1.00	10.0	4.	-1.	-1.	.28	-1.
3405	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.12	-1.	.100	1.8	5.3	.95	.60	2.	2.6	-1.	-1.	60.	-1.	0.	.04	-1.
3405	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.12	.12	.105	1.8	5.5	1.01	.64	2.	2.5	1.00	2.5	60.	-1.	-1.	.04	-1.
3406	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.12	-1.	.109	.9	4.3	.82	.48	11.	15.8	-1.	-1.	0.	-1.	-1.	.4	-1.
3406	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.12	.13	.115	.9	4.3	.82	.48	11.	15.0	1.00	15.0	0.	-1.	-1.	.40	-1.
3408	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.12	-1.	.094	1.7	4.3	.82	.56	4.	5.4	-1.	-1.	10.	5.	1.	.4	4.4
3408	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.12	.11	.095	1.8	4.3	.82	.58	4.	5.0	1.00	5.0	10.	-1.	-1.	.40	-1.
3409	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.12	-1.	.098	1.3	3.3	.63	.42	13.	19.7	-1.	-1.	7.	-1.	43.	1.6	8.0
3409	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.12	.115	.10	1.2	3.4	.64	.43	13.	19.0	1.12	21.5	7.	-1.	-1.	1.60	-1.
3410	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.12	-1.	.125	.3	4.3	.82	.42	8.	13.7	-1.	-1.	12.	-1.	5.	1.2	25.0
3410	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.12	.145	.13	.3	4.3	.82	.42	8.	12.0	1.00	12.0	12.	-1.	-1.	1.20	-1.
3412	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.14	-1.	.089	4.3	6.3	1.05	.85	9.	9.9	-1.	-1.	5.	-1.	0.	.34	2.1
3412	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.14	.11	.095	4.3	6.4	1.08	.87	9.	9.5	1.00	9.5	5.	-1.	-1.	.34	-1.
3413	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.14	-1.	.111	2.4	6.3	1.11	.72	8.	9.6	-1.	-1.	4.	-1.	0.	.36	2.6
3413	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.14	.13	.115	2.4	6.4	1.17	.77	8.	9.0	1.00	9.0	4.	-1.	-1.	.36	-1.
3414	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.14	-1.	.080	3.0	3.3	.61	.58	11.	14.6	-1.	-1.	5.	-1.	-1.	.53	2.5
3414	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.14	.095	.08	3.0	3.4	.63	.61	11.	13.5	1.00	13.5	5.	-1.	-1.	.53	-1.
3417	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.14	-1.	.093	2.5	4.0	.76	.61	6.	7.8	-1.	-1.	10.	4.	0.	.25	3.5
3417	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.14	.11	.095	2.4	4.0	.76	.61	6.	7.5	1.09	8.0	10.	-1.	-1.	.25	-1.
3419	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.14	-1.	.100	2.5	5.0	.95	.70	9.	10.9	-1.	-1.	20.	7.	0.	.15	15.0
3419	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.14	.12	.105	2.4	5.2	1.00	.72	9.	10.5	1.09	11.5	20.	-1.	-1.	.15	-1.
3421	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.14	-1.	.105	2.5	6.0	1.14	.79	12.	13.7	-1.	-1.	3.	-1.	2.	.35	2.3
3421	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.14	.125	.11	2.4	6.1	1.17	.81	12.	13.0	1.09	14.5	3.	-1.	-1.	.35	-1.
3424	7	0.	0.0	6.7	-1.	-1.	-1.	-1.	-1.	.12	-1.	.098	1.4	4.0	.76	.50	4.	5.7	-1.	-1.	10.	-1.	0.	.15	2.0
3424	8	0.	-1.	6.7	-1.	-1.	-1.	-1.	-1.	.12	.11	.10	1.5	4.0	.76	.52	4.	5.5	1.09	6.0	10.	-1.	-1.	.15	-1.
3501	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.20	-1.	.206	1.0	6.3	1.21	.68	10.	12.3	-1.	-1.	0.	-1.	0.	.60	-1.
3501	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.20	.22	.22	.9	6.4	1.23	.68	10.	12.0	1.09	13.0	0.	-1.	-1.	.60	-1.
3502	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.32	-1.	.314	.9	3.3	.63	.39	19.	29.6	-1.	-1.	4.	-1.	2.	.28	-1.
3502	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.32	.33	.325	.9	3.4	.64	.40	19.	28.5	1.12	31.5	4.	-1.	-1.	.28	-1.
3503	7	0.	0.7	7.4	-1.	-1.	-1.	-1.	-1.	.32	-1.	.350	.5	3.3	.63	.42	5.	8.1	-1.	-1.	3.	-1.	18.	.70	5.0
3503	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.32	.355	.35	.6	3.4	.64	.37	5.	8.0	1.00	8.0	5.	-1.	-1.	.70	-1.
3504	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.32	-1.	.292	1.3	3.3	.63	.42	7.	10.6	-1.	-1.	4.	-1.	2.	.28	2.5
3504	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.32	.305	.30	1.2	3.4	.64	.43	7.	10.0	1.00	10.0	4.	-1.	-1.	.28	-1.
3505	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.224	1.8	5.3	.95	.60	2.	2.6	-1.	-1.	60.	-1.	-1.	.04	-1.
3505	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.235	.23	1.8	5.5	1.01	.64	2.	2.5	1.00	2.5	60.	-1.	-1.	.04	-1.
3506	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.245	.9	4.3	.82	.48	11.	15.8	-1.	-1.	0.	-1.	-1.	.4	-1.
3506	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.255	.25	.9	4.3	.82	.48	11.	15.0	1.00	15.0	0.	-1.	-1.	.40	-1.
3507	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.206	2.2	5.3	.99	.68	20.	24.6	-1.	-1.	0.	-1.	-1.	.6	-1.
3507	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.22	.215	2.1	5.5	1.05	.72	20.	23.0	1.12	26.0	0.	-1.	-1.	.60	-1.
3508	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.210	1.7	4.3	.82	.56	4.	5.4	-1.	-1.	10.	5.	1.	.4	4.4

Table A 4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
3508	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.215	.215	1.8	4.3	.82	.58	4.	5.0	1.00	5.0	10.	-1.	-1.	.40	-1.
3509	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.219	1.3	3.3	.63	.42	13.	19.7	-1.	-1.	7.	-1.	43.	1.6	8.0
3509	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.23	.225	1.2	3.4	.64	.43	13.	19.0	1.12	21.5	7.	-1.	-1.	1.60	-1.
3510	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.281	.3	4.3	.82	.42	8.	12.1	-1.	-1.	12.	-1.	5.	1.2	25.0
3510	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.29	.29	.3	4.3	.82	.42	8.	12.0	1.00	12.0	12.	-1.	-1.	1.20	-1.
3511	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.241	1.3	7.3	1.39	.79	17.	18.4	-1.	-1.	17.	-1.	4.	.35	-1.
3511	9	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.26	.255	1.2	7.3	1.41	.80	17.	18.5	1.12	21.0	17.	-1.	-1.	.35	-1.
3512	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.172	4.3	6.3	1.05	.85	9.	9.9	-1.	-1.	5.	-1.	0.	.34	2.1
3512	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.185	.18	4.3	6.4	1.08	.87	9.	9.5	1.00	9.5	5.	-1.	-1.	.34	-1.
3513	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.214	2.4	6.3	1.11	.72	8.	9.6	-1.	-1.	4.	-1.	0.	.36	2.6
3513	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.225	.22	2.4	6.4	1.17	.77	8.	9.0	1.00	9.0	4.	-1.	-1.	.36	-1.
3514	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.28	-1.	.179	3.0	3.3	.61	.58	11.	14.6	-1.	-1.	5.	-1.	0.	.53	2.5
3514	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.28	.185	.185	3.0	3.4	.63	.61	11.	13.5	1.00	13.5	5.	-1.	-1.	.53	-1.
3515	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.28	-1.	.223	3.0	6.0	1.12	.82	23.	25.7	-1.	-1.	0.	-1.	0.	.41	2.5
3515	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.28	.235	.23	3.0	6.1	1.17	.67	23.	24.0	1.21	29.0	0.	-1.	-1.	.41	-1.
3516	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.202	2.5	6.0	1.14	.79	10.	11.4	-1.	-1.	10.	-1.	0.	.3	3.5
3516	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.215	.21	2.4	6.1	1.17	.82	10.	11.0	1.09	12.	10.	-1.	-1.	.30	-1.
3517	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.180	2.5	4.0	.76	.61	6.	7.8	-1.	-1.	10.	4.	0.	.25	3.6
3517	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.19	.185	2.4	4.0	.76	.61	6.	7.5	1.09	8.0	10.	-1.	-1.	.25	-1.
3518	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.208	2.5	7.0	1.33	.88	21.	22.6	-1.	-1.	5.	-1.	0.	.35	2.5
3518	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.225	.22	2.4	7.0	1.35	.89	21.	22.0	1.21	26.5	5.	-1.	-1.	.35	-1.
3519	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.193	2.5	5.0	.95	.70	9.	10.9	-1.	-1.	20.	7.	0.	.15	15.0
3519	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.205	.205	2.4	5.2	1.00	.72	9.	10.5	1.09	11.5	20.	-1.	-1.	.15	-1.
3520	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.187	2.5	4.5	.85	.65	10.	12.6	-1.	-1.	26.	8.	0.	.12	12.0
3520	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.20	.195	2.4	4.6	.88	.66	10.	12.0	1.09	13.0	26.	-1.	-1.	.12	-1.
3521	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.202	2.5	6.0	1.14	.79	12.	13.7	-1.	-1.	3.	-1.	2.	.35	2.3
3521	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.22	.215	2.4	6.1	1.17	.81	12.	13.0	1.09	14.5	3.	-1.	-1.	.35	-1.
3522	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.209	2.4	6.0	1.09	.73	15.	17.8	-1.	-1.	11.	4.	0.	.30	5.0
3522	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.22	.22	2.4	6.1	1.13	.77	15.	16.5	1.21	20.0	11.	-1.	-1.	.30	-1.
3523	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.24	-1.	.188	3.6	7.0	1.27	.93	17.	17.7	-1.	-1.	12.	3.	0.	.30	5.0
3523	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.24	.20	.195	3.7	7.0	1.30	.96	17.	17.0	1.21	20.5	12.	-1.	-1.	.30	-1.
3524	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.20	-1.	.183	1.4	4.0	.76	.50	4.	5.7	-1.	-1.	10.	-1.	0.	.15	2.0
3524	8	1.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.20	.185	.18	1.5	4.0	.76	.52	4.	5.5	1.09	6.0	10.	-1.	-1.	.15	-1.
3525	7	0.	0.0	7.4	-1.	-1.	-1.	-1.	-1.	.20	-1.	.195	1.4	6.0	1.14	.68	15.	18.5	-1.	-1.	10.	-1.	0.	.18	2.4
3525	8	0.	-1.	7.4	-1.	-1.	-1.	-1.	-1.	.20	.205	.20	1.5	6.1	1.17	.71	15.	17.5	1.21	21.0	10.	-1.	-1.	.18	-1.
3901	7	0.	0.0	6.6	-1.	-1.	-1.	-1.	-1.	.60	-1.	.449	1.8	4.0	.74	.52	31.	43.2	-1.	-1.	11.	-1.	-1.	.12	2.3
3901	8	0.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.78	.67	.575	1.8	3.7	.69	.51	28.	38.0	1.05	39.5	25.	-1.	-1.	.11	-1.
3902	7	0.	0.0	6.6	-1.	-1.	-1.	-1.	-1.	.60	-1.	.449	1.8	4.0	.74	.52	4.	5.6	-1.	-1.	25.	5.	-1.	.12	11.0
3902	8	0.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.78	.67	.575	1.8	3.7	.69	.51	1.	1.5	1.05	1.5	29.	-1.	-1.	.12	-1.
3903	7	0.	0.0	6.6	-1.	-1.	-1.	-1.	-1.	.60	-1.	.449	1.8	4.0	.74	.52	11.	15.3	-1.	-1.	19.	1.	-1.	.10	2.4
3903	8	0.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.78	.695	.595	1.8	4.3	.81	.57	13.	16.5	1.05	17.5	37.	-1.	-1.	.09	-1.

Table A.4 (Continued)

CASE	BCAT	LIQ1	LIQ2	M	H	EP	HY	DER	DUR	A	CSR	CSRN	ZW	ZL	SIGT	SIGE	N	N1	CE	N160	FC	CC	GC	D50	UC
3904	7	1.	0.7	6.6	-1.	-1.	-1.	-1.	-1.	.20	-1.	.206	.2	2.0	.38	.20	3.	5.7	-1.	-1.	66.	15.	-1.	.04	-1.
3904	8	1.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.24	.265	.23	.3	1.8	.36	.21	3.	5.5	.79	4.5	80.	-1.	-1.	.04	-1.
3905	7	1.	0.7	6.6	-1.	-1.	-1.	-1.	-1.	.20	-1.	.209	.2	5.0	.95	.47	7.	10.1	-1.	-1.	34.	4.	-1.	.09	-1.
3905	8	1.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.24	.28	.24	.3	4.3	.85	.45	11.	15.0	1.05	16.0	18.	-1.	-1.	.15	-1.
3918	8	1.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.20	.16	.135	2.1	3.4	.62	.50	2.	2.5	1.05	3.0	75.	-1.	-1.	.047	-1.
3919	8	0.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.20	.13	.115	2.1	2.3	.41	.39	11.	16.5	.79	13.0	30.	-1.	-1.	.100	-1.
3925	8	1.	-1.	6.6	-1.	-1.	-1.	-1.	-1.	.51	.39	.335	1.5	2.1	.39	.33	3.	5.0	.79	4.0	31.	-1.	-1.	.105	-1.
4101	7	0.	0.0	6.1	-1.	-1.	-1.	-1.	-1.	.10	-1.	.086	1.0	6.0	1.08	.58	5.	6.6	-1.	-1.	13.	-1.	0.	.18	-1.
4101	8	0.	-1.	6.1	-1.	-1.	-1.	-1.	-1.	.10	.12	.095	.9	6.1	1.10	.59	5.	6.5	1.09	7.0	13.	-1.	-1.	.18	-1.
4102	7	1.	0.7	6.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.143	-1.	-1.	-1.	-1.	-1.	6.6	-1.	-1.	13.	-1.	0.	.18	-1.
4102	8	0.	0.5	6.1	-1.	-1.	-1.	-1.	-1.	.145	.17	.135	.9	6.1	1.10	.59	5.	6.5	1.09	7.0	13.	-1.	-1.	.18	-1.
4103	7	0.	0.0	6.1	-1.	-1.	-1.	-1.	-1.	.10	-1.	.094	1.0	14.0	2.52	1.08	4.	3.8	-1.	-1.	27.	-1.	0.	.17	-1.
4103	8	0.	-1.	6.1	-1.	-1.	-1.	-1.	-1.	.10	.105	.085	.9	14.3	2.59	1.25	4.	3.5	1.09	4.0	27.	-1.	-1.	.17	-1.
4104	7	1.	0.7	6.1	-1.	-1.	-1.	-1.	-1.	-1.	-1.	.165	-1.	-1.	-1.	1.08	4.	3.8	-1.	-1.	27.	-1.	0.	.17	-1.
4104	8	1.	0.5	6.1	-1.	-1.	-1.	-1.	-1.	.145	.155	.12	.9	14.3	2.59	1.25	4.	3.5	1.09	4.0	27.	-1.	-1.	.17	-1.
4204	8	0.	-1.	5.6	-1.	-1.	-1.	-1.	-1.	.21	.235	.165	.3	1.8	.38	.21	3.	5.5	.79	4.5	80.	-1.	-1.	.04	-1.
4205	8	0.	-1.	5.6	-1.	-1.	-1.	-1.	-1.	.21	.245	.175	.3	4.3	.85	.45	11.	15.0	1.05	16.0	18.	-1.	-1.	.15	-1.
4218	8	1.	-1.	5.6	-1.	-1.	-1.	-1.	-1.	.20	.16	.11	2.1	3.4	.62	.50	2.	2.5	1.05	3.0	75.	-1.	-1.	.047	-1.
4219	8	0.	-1.	5.6	-1.	-1.	-1.	-1.	-1.	.20	.13	.095	2.1	2.3	.41	.39	11.	16.5	.79	13.0	30.	-1.	-1.	.100	-1.
4225	8	1.	-1.	5.6	-1.	-1.	-1.	-1.	-1.	.26	.28	.20	.9	4.9	.98	.56	9.	11.0	1.05	11.5	40.	-1.	-1.	.09	-1.
4226	8	1.	-1.	5.6	-1.	-1.	-1.	-1.	-1.	.32	.26	.185	2.7	4.3	.79	.64	5.	6.5	1.05	6.5	92.	-1.	-1.	.045	-1.
4227	8	0.	-1.	5.6	-1.	-1.	-1.	-1.	-1.	.09	.07	.05	1.5	2.1	.39	.33	3.	5.0	.79	4.0	31.	-1.	-1.	.105	-1.

Table A.5

CASE SOURCE REFERENCE GUIDE

Case No.	Original Source Catalog(s)	Primary Source(s)	Supplementary Source(s)
0101 to 0102	2	Koizumi (1966)	Seed and Idriss (1967) Kawasumi (1968)
0201 to 0202	2	Koizumi (1966)	Seed and Idriss (1967) Kawasumi (1968)
0301 to 0305	2	Kishida (1969)	Tokimatsu (1984)
0306 to 0307	6	Ishihara (1974)	Kishida (1969)
0401 to 0405	6	Kuribayashi et al. (1977)	Kuribayashi and Tatsuoka (1975)
0501 to 0506	3,6	*D&M (1976) Youd and Hoose (1976)	Clough and Chameau (1983) Hanks and Kanamori (1979) Lawson et al. (1908) NOAA (1972) Real et al (1978) Youd and Hoose (1978)
0507 to 0508	0	Clough and Chameau (1983)	NOAA (1972) Real et al. (1978)
0601 to 0602	6	Ishihara (1974)	Kuribayashi and Tatsuoka (1975) Kuribayashi et al. (1977)
0701 to 0707	7	*Kodera (1964)	
0708 to 0712	6	Kuribayashi et al. (1977)	Kuribayashi and Tatsuoka (1975)

Table A.5
(Continued)

Case No.	Original Source Catalog(s)	Primary Source(s)	Supplementary Source(s)
0713 to 0714	6	Ishihara (1974)	
0801	2	Seed et al (1969)	
0901 to 0905	6	Kuribayashi et al. (1977)	Kuribayashi and Tatsuoka (1975)
1001 to 1004	3	*D&M (1976)	D&M (1977) Knuppel (1974) Pyke et al. (1978) Real et al. (1978)
1005	0	Pyke et al. (1978)	D&M, (1975) Knuppel (1974)
1101 to 1103	2	Ross (1968)	
1201 to 1204	2,7	Kishida (1969)	Tokimatsu (1984)
1301 to 1306	2,7	Kishida (1969)	
1401	3	*D&M (1976)	Clough and Chameau (1983) Brazee and Cloud (1959) Yegian (1984)
1501 to 1510	3	*D&M (1976)	Clough and Chameau (1983) Brazee and Cloud (1959) Yegian (1984)

Table A.5
(Continued)

Case No.	Original Source Catalog(s)	Primary Source(s)	Supplementary Source(s)
1511 to 1532	3	*W.C.C.(1976)	Clough and Chameau (1983) Brazee and Cloud (1959) Yegian (1984)
1533 to 1534	3	*D&M (1976)	Clough and Chameau (1983) Brazee and Cloud (1959) Yegian (1934)
1535 to 1536	0	Clough and Chameau (1983)	Brazee and Cloud (1959)
1601 to 1603	2	*Lee (1970)	Housner (1963) Duke and Leeds (1963) Seed (1984)
1604 to 1605	1	*Poblete (1967)	Housner (1963) Blume (1963)
1701 to 1704	2	Ross et al. (1969)	Ross (1968)
1705	2	Coulter and Migliaccio (1966)	
1706	1	Ross et al. (1969)	Ross (1968)
1801	2	Koizumi (1966)	Seed and Idriss (1967)
1802	2	Kishida (1966)	
1803 to 1807	1,2	Koizumi (1966)	Seed and Idriss (1967)
1808	3	Kishida (1966)	
1809 to 1810	4	Kawakami and Asada (1966)	

Table A.5
(Continued)

Case No.	Original Source Catalog(s)	Primary Source(s)	Supplementary Source(s)
1811 to 1812	4	Kishida (1966)	
1813	4	Ohsaki (1966)	
1814 to 1815	7	Koizumi (1966)	
1816 to 1819	7	Kishida (1966) Ohsaki (1966)	
1820 to 1821	7	*Takada et al. (1965)	
1822 to 1823	7	Ishihara et al. (1979)	
1901	3	*D&M (1976)	NOAA (1972) Real et al. (1978)
2001	2	*Cluff (1973)	
2101 to 2102	4	Yamanouchi et al. (1970)	
2201	1,2	Yoshimi and Akagi (1963) Kishida (1970)	Yoshimi (1970)
2202 to 2210	1,2,4	Ohsaki (1970)	Yoshimi (1970)
2211	4	Yoshimi and Akagi (1968)	Yoshimi (1970)
2212	4	Yoshimi (1970)	

Table A.5
(Continued)

Case No.	Original Source Catalog(s)	Primary Source(s)	Supplementary Source(s)
2301 to 2306	4	*ERI (1969)	
2401 to 2402	0	Clough and Chameau (1983)	Steinbrugge et al. (1970) Real et al. (1978)
2501	4	Whitman et al. (1974)	Tezcan and Ipek (1973) Tezcan et al. (1977)
2601	2	Seed (1973)	NOAA (1973)
2602	2	Dixon and Burke (1973)	NOAA (1973)
2603	7	Lee et al. (1975) Seed, Idriss et al. (1975) Seed, Lee et al. (1975)	NOAA (1973)
2604 to 2611	3	WCC (1976)	NOAA (1973)
2701	3,4	Sato, et al. (1973)	
2801 2802 2804 2806	5	None cited in base catalog	Fu Shengcong and Tatsuoka (1983, 1985)
2803 2805 2807 2809 2810 2811 2812	5,8	None cited in base catalog No. 5 Fu Shengcong and Tatsuoka (1983, 1985)	
2813 to 2815	8	Fu Shengcong and Tatsuoka (1983, 1985)	

Table A.5
(Continued)

Case No.	Original Source Catalog(s)	Primary Source(s)	Supplementary Source(s)
2901 to 2906	7,8	Seed, et al. (1981)	
3001 to 3002	5	None cited in base catalog, No. 5 Fu Shengcong and Tatsuoka (1983, 1985)	Ye and Liu (1979)
3303 to 3005	7	Zhou (1981)	Ye and Liu (1979)
3101 to 3102	0	Ishihara and Perlea (1984)	
3201 to 3206	0	Idriss et al (1979)	Youd and Keefer (1984) Brogan and Slemmons (1985) Slemmons and Brogan (1985)
3207 to 3217	0	Youd (1984)	Youd and Keefer (1984)
3301	7	Okusa et al. (1980)	Ishihara (1984)
3401 to 3413	7	*Iwasaki, Kawashima and Tokida (1978)	Iwasaki, Tokida and Tatsuoka (1981) Tokimatsu (1984)
3414 to 3415	1	*Iwasaki et al. (1981)	Iwasaki, Tokida and Tatsuoka (1981) Tokimatsu (1984)
3416 to 3423	7	*Tsuchida et al. (1979)	Iwasaki, Tokida and Tatsuoka (1981) Tokimatsu (1984)
3424 to 3425	7	Ishihara et al. (1980)	Tokimatsu (1984)

Table A.5
(Continued)

Case No.	Original Source Catalog(s)	Primary Source(s)	Supplementary Source(s)
3501 to 3513	7	*Iwasaki, Kawashima and Tokida (1978)	Iwasaki, Tokida and Tatsuoka (1981) Tokimatsu (1974)
3514 to 3515	7	*Iwasaki, et al. (1981)	Iwasaki, Tokida and Tatsuoka (1981) Tokimatsu (1984)
3516 to 3523	7	*Tsuchida et al. (1979)	Iwasaki, Tokida and Tatsuoka (1981) Tokimatsu (1984)
3524 to 3527	7	Ishihara et al. (1980)	Tokimatsu (1984)
3601 to 3602	0	Gazetas and Botsis (1981)	Gazetas (1984)
3701 to 3702	0	Jaime, Montanez, and Romo (1981)	Jaime, Romo and Montanez (1981)
3801	0	Talaganov et al. (1980, 1981)	ISC (1979) EERI (1980)
3901 to 3924	0,7	Bennett et al. (1981) Youd and Bennett (1983) Bennett et al. (1984)	Diaz-Rodriquez (1984) Chavez et al. (1982)
4001	0,8	Diaz-Rodriguez (1984)	
4101	7	Ishihara et al. (1981)	

Table A.5
(Continued)

Case No.	Original Source Catalog(s)	Primary Source(s)	Supplementary Source(s)
4201 to 4224	0,8	Bennett et al. (1984) Youd (1984)	Youd and Wieczorek (1984) Stokoe (1984) Bierschwale (1984)

*Indicates reference not available to the writer.

Table A.6

MAGNITUDE SCALES FOR TABULATED EARTHQUAKES

Catalog Code	Earthquake	Date	Tabulated Magnitude	Magnitude Scale
01	Niigata Sado Island	1802 12/1	6.6	M_k
02	Niigata Koshigun	1877 7/22	6.1	M_k
03	Mino-Owari	1891 10/28	7.9	JMA*
04	Tokyo	1894 6/20	7.5	JMA*
05	San Francisco	1906	8.3	M_L†
06	Gono	1909 8/14	6.9	M_k
07	Kanto	1923 9/1	7.9	JMA*
08	Santa Barbara	1925	6.3	M_L†
09	Nishi-Saitama	1931 9/21	7.0	JMA
10	Long Beach	1933 3/10	6.3	Unspecified M_L ?
11	El Centro	1940	7.0	M_L
12	Tonankai	1944 12/7	8.0	JMA*
13	Fukui	1948	7.3	JMA*
14	San Francisco	1955	5.4	M_L
15	San Francisco Daly City	1957 3/22	5.3	M_L
16	Chile[1]	1960 5/22	8.4 7.5	M_L

Table A.6 (Continued)

Catalog Code	Earthquake	Date	Tabulated Magnitude	Magnitude Scale
17	Alaska	1964 3/27	8.3	M_S
18	Niigata	1964 6/16	7.5	JMA
19	San Francisco Concord Bay	1965	4.9	M_L
20	Caracas	1967	6.3	Unspecified M_L?
22	Tokachi-Oki	1968 5/16	7.9	M_L
23	Saitama	1968 7/1	6.1	JMA?
24	Santa Rosa	1969	5.7	M_L
25	Gediz, Turkey	1970	7.1	M_L
26	San Francisco	1971 2/9	6.6	M_L
27	Yokohama	1972	7.3	JMA
28	Haicheng	1975 2/4	7.3	M_S
29	Guatemala	1976 2/4	7.5	M_L
30	Tangshan[2]	1977 7/28	7.8 7.6	M_S Unspecified
32	San Juan Argentina	1977 11/3	7.4	M_L
33	Izu	1978	7.0	JMA?
34	Miyagiken-Oki-1	1978 2/20	6.7	JMA
35	Miyagiken-Oki-2	1978 6/12	7.4	JMA

Table A.6 (Continued)

36	Thessaloniki	1978 6/20	6.5	Unspecified M_S?
37	Guerrero	1979 3/14	7.6	M_L
38	Montenegro	1979	6.9	M_S
39	Imperial Valley	1979 10/15	6.6	M_L
40	Mexicali Valley	1980 9/25	6.7	M_L
41	Mid-Chiba	1980 9/25	6.1	JMA?
42	Westmoreland	1981 4/26	5.6	M_L

Key: M_L = Gutenberg-Richter (Local) Magnitude

M_S = Shear Wave Magnitude

M_K = Kawasumi Magnitude

JMA = Japan Meteorological Agency Magnitude

(*) indicates JMA magnitude estimated from Kawasumi scale by Tokimatsu and Yoshimi (1983)

(†) indicates M_L not measured, but estimated through intensity or some other means.

(?) indicates some uncertainty regarding scale, usually because it is unspecified.

Notes: 1) Two magnitudes specified for Chile earthquake. M_L = 7.5 was a foreshock, but epicenter was closer to case study than M_L = 8.4 shock.

2) That there are two magnitudes specified for the Tangshan Earthquake (each for different case studies) was an error not noted until after completion of analyses. However, error will not significantly affect results.

BIBLIOGRAPHY FOR APPENDIX A

LIQUEFACTION CATALOG SOURCE REFERENCES

†Arango, I., Tokimatsu, K., and Beratan, L. (1983), "A Survey of the Ground Liquefaction Resulting from the Nihon-Kai-Chubu Earthquake," Report submitted to the National Science Foundation, Washington, D.C., 1983.

Bennett, M.J., McLaughlin, Sarmiento, J.S. and Youd, T.L. (1984), "Geotechnical Investigation of Liquefaction Sites," Open File Report 84-252, U.S. Geological Survey, Menlo Park, California, 1984.

Bennett, M.J., Youd, T.L., Harp, E.L. and Weiczorek, G.F. (1981), "Subsurface Investigation of Liquefaction, Imperial Valley Earthquake, California, October 15, 1979," Open-File Report 81-502, U.S.Geological Survey, Menlo Park, California, 1981.

Biershwale, J.G. (1984), "Analytical Evaluation of Liquefaction Potential of Sands Subjected to the 1981 Westmorland Earthquake," Geotechnical Engineering Report GR-84-15, Civil Engineering Department, University of Texas, Austin, Texas, 1984.

Blume, J.A. (1963), "A Structural-Dynamic Analysis of Steel Plant Structures Subjected to the May 1960 Chilean Earthquakes," Bulletin of the Seismological Society of America, Vol. 53, No. 2, Feb. 1963, pp. 439-480.

Brazee, R.J. and Cloud, W.K. (1959), United States Earthquakes - 1957, U.S. Dept. of Commerce, Coast and Geodetic Survey, Washington, D.C., 1959.

Brown, G.E. and Slemmons (1985), "Liquefaction Effects and Ground Failures from the November 23, 1977 San Juan, Argentina Earthquake," U.S. Geological Survey Publication (In Press).

†Carillo-Gil, A. (1981), "Comparative Studies of Soil Liquefaction Potential During the 1970 Peru Earthquake," Proceedings, International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 2, pp. 687-689.

Clough, G.W. and Chameau J.L. (1983), "Seismic Response of San Francisco Waterfront Fills," Jour. of Geotechnical Engineering, ASCE, Vol. 109, No. 4, April 1983, pp. 491-506.

Chavez, D., Gonzales, J., Reyes, A., et al. (1982), "Main-Shock Location and Magnitude Determination Using Combined U.S. and Mexican Data," The Imperial Valley Earthquake, U.S. Geological Survey Prof. Paper 1254, U.S. Govt. Printing Office, Washington, D.C., 1982, pp. 51-54.

*Cluff, L.S. (1973), Personal Communication to Seed, et al., 1973.

Coulter, H.W. and Migliaccio, R.R. (1966), "Effects of the Earthquake of March 27, 1964 at Valdez, Alaska," Professional Paper 542-C, U.S. Geological Survey, 1966.

*D & M (1976), Data from Dames and Moore files, reported by Yegian, 1976.

D & M (1977), "Detailed Geotechnical Investigation Proposed Crude Oil Tank Farm, Pier J, Port of Long Beach," Dames & Moore Report to the Standard Oil Company, Ohio, Apr. 1977.

D & M (1975), "Offshore Soils Investigation, Los Angeles Harbor, LNG Ship Terminal," Dames & Moore Report to Western LNG Terminal Company, Sept. 1975.

Davis, R.O. and Berrill, J.B. (1981), "Assessment of Liquefaction Potential Based on Seismic Energy Dissipation," Proc. Int'l Conf. on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics," St. Louis, Missouri, 1981, Vol. 1, pp. 187-190.

†Diaz De Cossio, R. (1960), "Foundation Failures During the Coatzacoalas (Mexico) Earthquake of 26 August 1959," Proc. 2nd World Conf. on Earthquake Engr., Tokyo, Vol. 1, pp. 473-486.

Diaz-Rodriguez, J.A. (1984), "Liquefaction in the Mexicali Valley During the Earthquake of June 9, 1980," Proc. 8th World Conf. on Earthquake Engineering, San Francisco. 1984. Vol. 3, pp. 223-230.

Dixon, S.J. and Burke, J.W. (1973), "Liquefaction Case History," Jour. of the Soil Mech. and Foundation Div., ASCE, Vol. 99, No. SM 11, Nov. 1973, pp. 921-937.

†Dobry, R. and Alvarez, L. (1967), "Seismic Failures of Chilean Talings Dams," Jour. of the Soil Mechanics and Foundations Div., ASCE, Vol. 93, No. SM6, Nov. 1967, pp. 237-260.

Duke, C.M. and Leeds, D.J. (1963), "Response of Soils, Foundations, and Earth Structures to the Chilean Earthquakes of 1960," Bull. Seismological Soc. of America, Vol. 53, No. 2. Feb. 1963, pp. 309-357.

†EERI (1978), "Reconnaissance Report, Miyagi-Ken-Oki, Japan Earthquake, June 12, 1978," Earthquake Engineering Research Institute, Berkeley, Calif., Dec. 1978.

EERI (1980), "Reconnaissance Report, Montenegro, Yugoslavia Earthquake, April 15, 1979," Earthquake Engineering Research Institute, Nov. 1980.

†Egan, J.A. (1984), Personal Communication with S. Liao, 1984.

Egan, J.A., Moriwaki, Y. and Moses, T.L. (1984), "Site Characterization for Seismic Hazards Evaluations," Presented at the EERI Symposium on on Earthquake Engineering in Alaska, Anchorage, Alaska, June 1984.

*ERI (1969), "Strong Motion Earthquake Records in Japan, Vol. 7," Earthquake Research Institute, University of Tokyo, March 1969.

Fu Shengcong and Tatsuoka, F. (1983), "Soil Liquefaction During Haicheng and Tangshan Earthquakes in China," Report of Japan-China Cooperative Research on Engineering Lessons from Recent Chinese Earthquakes, Including the 1976 Tangshan Earthquake. (Part 1), Edited by C. Tamura, T. Katayama and F. Tatsuoka, Institute of Industrial Science, Univ. of Tokyo, Nov. 1983, pp. 129-286.

Fu Shengcong and Tatsuoka, F. (1984), "Soil Liquefaction During Haicheng and Tangshan Earthquake in China; A Review," Soils and Foundations, Vol. 24, No. 4, Dec. 1984, pp. 12-29.

Gazetas, G. (1984),Personal Communication with S. Liao, Oct. 1984.

Gazetas, G. and Botsis, J. (1981), "Local Soil Effects and Liquefaction in the 1978 Thessaloniki Earthquakes," Proc. Int'l Conf. on Recent Advances in Geotech. Earthquake Engr. and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 3, pp. 1025-1213

Housner, G.W. (1963), "An Engineering Report on the Chilean Earthquakes of May 1960, Bull. Seismological Soc. of America, Vol. 53, No. 2, Feb. 1963, pp. 219-223.

Idriss, I.M., Arango, I. and Brogan, G. (1979), "Study of Liquefaction in November 23, 1977 Earthquake, San Juan Province, Argentina," Report to U.S. Geological Survey, Menlo Park, Ca. by Woodward-Clyde Consultants, San Francisco,CA., 1979.

ISC (1979), Bulletin of the International Seismologic Center, Newbury, Berkshire, England, 1979.

Ishihara, K. (1974), "Liquefaction of Subsurface Soils During Earthquakes," Technocrat, Vol. 7, No. 5, 1974, pp. 81-98.

Ishihara, K. (1984), "Post-Earthquake Failure of a Tailings Dam due to Liquefaction of the Pond Deposit," Proc. Int'l Conf. on Case Histories in Geotechnical Engineering," St. Louis, Missouri, May 1984, Vol. 3, pp. 1129-1143.

Ishihara, K., Kawase, Y., and Nakajima (1980), "Liquefaction Characteristics of Sand Deposits at an Oil Tank Site During the 1978 Miyagiken-Oki Earthquake," Soils and Foundations, Vol. 20, No. 2, June 1980, pp. 97-111.

Ishihara, K. and Perlea, V. (1984), "Liquefaction-Associated Ground Damage During the Vrancea Earthquake of March 4, 1977," Soils and Foundations, Vol. 24, No. 1, March 1984, pp. 90-112.

Ishihara, K. and Koga, Y. (1981), "Case Studies of Liquefaction in the 1964 Niigata Earthquake, Soils and Foundations, Vol. 21, No. 3, Sept. 1981, pp. 35-52.

Ishihara, K., Shimizu, K. and Yamada, Y. (1981), "Pore Water Pressures Measured in Sand Deposits During an Earthquake," Soils and Foundations, Vol. 21, No. 4, Dec. 1981, pp. 85-100.

Ishihara, K., Silver, M.L. and Kitagawa, H. (1979), "Cyclic strength of Undisturbed Sands Obtained by a Piston Sampler," Soils and Foundations, Vol. 19, No. 3, 1979, pp. 61-76.

†Ishihara, K. and Towhata, I. (1982), "Dynamic Response Analysis of Level Ground Based on the Effective Stress Method," Chapter 7, Soil Mechanics - Transient and Cyclic Loads, Edited by N. Pande and O.C. Zienkiewicz, John Wiley and Sons, 1982, pp. 133-172.

*Iwasaki, et. al. (1981), "Studies on Aseismic Stability of River Dykes Considering Ground Liquefaction," Proc. 16th Annual Meeting, JSSMFE, 1981, pp. 637-640 (in Japanese).

*Iwasaki, T., Kawashima, K. and Tokida, K. (1978), "Report of the Miyagiken-Oki Earthquake of June, 1978, Public Works Research Institute, Ministry of Construction, Report No. 1422, 1978 (in Japanese).

Iwasaki, T. and Tokida, K. (1980), "Studies on Soil Liquefaction Observed During the Miyagi-Ken Oki Earthquake of June 12, 1978," Proc. 7th World Conf. on Earthquake Engineering, Istanbul, Vol. 3, pp. 195-199.

Iwasaki, T., Tokida, K. and Tatsuoka, F. (1981), "Soil Liquefaction Potential Evaluation with Use of the Simplified Procedure," Proc. Int'l Conf. on Recent Advances in Geotech. Earthquake Engr. and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 1, pp. 209-214.

Jaime, A., Montanez, L. and Romo, M.P. (1981) "Liquefaction of Enmedio Island Soil Deposits," Proc. Int'l Conf. on Recent Advances in Geotechnical Earthquake Engr. and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 1, pp. 529-534.

Jaime, A., Romo, M.P. and Montanez, L. (1981) "Observed and Predicted Liquefaction of a Sand Stratum," Proc. Int'l Conf. on Recent Advances in Geotech. Earthquake Engr. and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 1, pp. 505-510.

Kawakami, H. (1968), "Historical Earthquakes in the Disturbed Area and Vicinity," General Report on the Niigata Earthquake of 1964, edited by H. Kawasumi, Electrical Engineering College Press, University of Tokyo, 1968, pp. 33-46.

Kawasumi, F. and Asada, A. (1966), "Damage to the Ground and Earth Structures by the Niigata Earthquake of June 16, 1964," Soils and Foundations, Vol. 6, No. 1, Jan. 1966, pp. 14-30.

Kishida, H. (1966), "Damage to Reinforced Concrete Buildings in Niigata City with Special Reference to Foundation Engineering," Soils and Foundations, Vol. 6, No. 1, 1966, pp. 71-88.

Kishida, H. (1969), "Characteristics of Liquefied Sands During Mino-Owari, Tohnankai and Fukui Earthquakes," Soils and Foundations, Vol. 9, No. 1, 1969, pp. 75-92.

Kishida, H. (1970), "Characteristics of Liquefaction of Level Sandy Ground During the Tokachioki Earthquake," Soils and Foundations, Vol. 10, No. 2, 1970, pp. 103-111.

Knuppel, L.A. (1974), "Liquefaction Potential of Proposed Fills, Los Angeles Harbor," M.S. Thesis, Univ. of Calif. Los Angeles, 1974.

†Kobayashi, H., Seo, K. Midorikawa, S., Yoshimi, Y., Tohno, I., Tokimatsu, K., Katayama, T. and Shibata, H., "A Report on the Miyagiken-Oki, Japan Earthquakes of June 12, 1978," Proc. 2nd Int'l Conf. on Microzonation, San Francisco, Nov.-Dec., 1978, pp. 587-611.

*Kodera, J. (1964), "Earthquake Damage and the Ground of Pier Foundations, Part 1," Tsuchi-to-Kiso, Vol. 12, No. 3, 1964, pp. 11-18 (in Japanese).

Koizumi, Y. (1966), "Changes in Density of Sand Subsoil Caused by the Niigata Earthquake," Soils and Foundations, Vol. 6, No. 2, 1966, pp. 38-44.

Kuribayashi, E. and Tatsuoka, T. (1975), "Brief Review of Liquefaction during Earthquakes in Japan," Soils and Foundations, Vol. 15, No. 4, Dec. 1975, pp. 81-92.

Kuribayashi, E., Iwasaki, T. and Tatsuoka, F. (1977), Proceedings 6th World Conference on Earthquake Engineering, New Delhi, 1977, Vol. 3, pp. 2448-2454.

Lawson, A.C., Gilbert, G.K., Reid, J.C. et al. (1908), "The California Earthquake of April 18, 1906," Carnegie Institute, Publication 87, Washington, D.C., 1908.

*Lee, K.L. (1970), Private Communication to H.B. Seed, Et. al., 1970.

Lee, K.L., Seed, H.B., Idriss, I.M. and Makdisi, F.I. (1975), "Properties of Soil in the San Fernando Hydraulic Fill Dams," Jour. Geotech Engr. Div., ASCE, Vol. 101, No. GT8, Aug. 1975, pp. 801-821.

†Moriwaki, Y., Idriss, I.M., Egan, J. and Moses, T.L. (1984), "A Reevaluation of the 1964 Fourth Avenue Slide," Presented at the EERI Symposium on Earthquake Engineering in Alaska, Anchorage, Alaska, June 1984.

†NBS (1980), "An Investigation of the Miyagiken-Oki, Japan, Earthquake of June 12, 1978," NBS Special Publication 592, U.S. Dept. of Commerce, National Bureau of Standards, Washington, D.C. Oct. 1980.

NOAA (1972), "A Study of Earthquake Losses in the San Francisco Bay Area Data and Analysis," U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, 1972.

NOAA (1973), "San Fernando, California Earthquake of February 9, 1971," National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, Washington, D.C., 1973 (3 volumes).

Ohaski, Y. (1966), "Niigata Earthquakes, 1964 Building Damage and Soil Condition," Soils and Foundations, Vol. 6, No. 2, 1966, pp. 14-37.

Ohaski, Y. (1970), "Effects of Sand Compaction on Liquefaction During the Tokachioki Earthquake," Soils and Foundations, Vol. 10, No. 2, June 1970, pp. 112-128.

Okusa, S., Anma, S. and Maikuma, H. (1980) "Liquefaction of of Mine Tailings in the 1978 Izu-Ohsima-Kinkai Earthquake, Central Japan," Proc. 7th World Conf. on Earthquake Engr., Istanbul, Vol. 3, pp. 89-96.

*Poblete, M. (1967), "El Subsuelo Del Centro De Concepcion En Relacion Con El Diseno Anti Sismico," Thesis presented to Universidad De Chile, Santiago Chile, in partical fullfillment of the requirements for the Degree in Civil Engineering, 1967.

Pyke, R.M., Knuppel, L.A. and Lee, K.L. (1978), "Liquefaction Potential of Hydraulic Fills," Jour. Geotech. Engr. Div., ASCE, Vol. 104, No. GT11, Nov. 1978, pp. 1335-1354.

Real, C.R., Toppozada, Parke, D.L. (1978), "Earthquake Catalog of California, January 1, 1900 - December 31, 1974," Special Publication 52, California Division of Mines and Geology, Sacramento, CA., 1978.

Ross, G.A. (1968), "Case Studies of Soil Stability Problems Resulting from Earthquakes," Ph.D. Dissertation, Univ. of California, Berkeley, 1968.

Ross, G.A. (1984), Personal Communication with S. Liao, Nov. 1984.

Ross, G.A., Seed, H.B., Migliaccio, R.R. (1969) "Bridge Foundation behavior in Alaska Earthquake," Journal of the Soil Mechanics and Foundations Division, ASCE, Vol. 95, No. SM4, July 1969, pp. 1007-1037.

Sato, N., Katayama, T. and Kubo, K. (1973), "Earthquake Observations at a 35,000 kl LNG Tank," Bull., Earthquake Resistant Structure Research Center, Institute of Industrial Science, Univ. of Tokyo, No. 7, Dec. 1973, pp. 17-35.

Seed, H.B. (1973), Personal Files of H.B. Seed, 1973.

Seed, H.B., Arango, I. and Chan, C. (1975), "Evaluation of Soil Liquefaction Potantial During Earthquakes," Report No. EERC 75-28, Univ. of California, Berkeley, Oct. 1975.

Seed, H.B., Arango, I., Chan, C.K., Gomez-Masso, A. and Grant-Ascoli, R. (1981), "Earthquake-Induced Liquefaction Near Lake Amatitlan, Guatemala," Jour. of the Geotech. Engr. Div., ASCE, Vol. 107, No. GT4, Apr. 1981, pp. 501-518.

Seed, H.B. and Idriss, I.M. (1967), "Analysis of Soil Liquefaction: Niigata Earthquake," Jour. of the Soil Mechanics and Foundations Division, ASCE, Vol. 97, No. SM6, Nov. 1967, pp. 105-134.

Seed, H.B., Idriss, I.M., Lee, K.L. and Makdisi, F.I. (1975), "Dynamic Analysis of the Slide in the Lower San Fernando Dam During the Earthquake of February 9, 1971," Jour. of the Geotech. Engr. Div., ASCE, Vol. 101, No. GT9, pp. 889-911.

Seed, H.B., Lee, K.L. and Idriss, I.M. (1969), "An Analysis of the Sheffield Dam Failure," Jour. of the Soil Mechanics and Foundations Division, ASCE, Vol. 95, No. SM6, Nov. 1969, pp. 1453-1490.

Seed, H.B., Lee, K.L., Idriss, I.M. and Makdisi, F.I. (1975), "The Slides in the San Fernando Dams During the Earthquake of February 9, 1971," Jour. of the Geotech. Engr. Division, ASCE, Vol. 101, No. GT7, July 1975, pp. 651-688.

†Seed, H.B. and Wilson, S.D. (1967), "The Turnagin Heights Landslide, Anchorage, Alaska,", Jour. Soil Mech. and Found. Div., ASCE, Vol. 93, No. SM4, July 1967, pp. 325-353.

Slemmons, D.B. and Brogan, G.E. (1985), "Tectonic Setting and Primary and Geological Effects of the November 23, 1977 Earthquake, San Juan Province, Argentina," U.S. Geological Survey Publication (in Press).

Steinburgge, K.V., Cloud, W.K. and Scott, N.H. (1970), "The Santa Rosa, California Earthquakes of October 1, 1969," U.S. Dept. of Commerce, Coast and Geodetic Survey, Rockville, Maryland, U.S. Govt Printing Office, Washington, D.C. 1970.

Stokoe, K.H. II (1984), "Investigation of Areas of Recurring Liquefaction in the Vicinity of the New and Alamo Rivers, Imperial Valley, California," Summaries of Technical Reports, Volume XVIII Prepared by Participants in National Earthquake Hazards Reduction Program, Open-File Report 84-628, U.S. Geological Survey, Menlo Park, CA, 1984.

*Takada, T. et al., (1965), "A Report of Niigata Earthquake-Part 5, Damage of Bridges," Report No. 125-5, Public Works Research Institute, Ministry of Construction, 1965 (in Japanese).

Talaganov, K., Petrovski, J. and Mihailov, V. (1981), "Soil Liquefaction Seismic Risk Analysis Based on Post 1979 Earthquake Observations in Montenegro," Proceedings, Int. Conf. on Recent Advances in Geot. Earthquake Engr. and Soil Dynamics, St. Louis, Missouri, 1981, Vol. 2, pp. 691-697.

Talaganov, K., Mihailov, V. and Bogoevski, T. (1980), "Analysis of Soil Liquefaction during 1979 Monte Negro Earthquake," Proc. 7th World Conf. on Earthquake Engineering, Istanbul, 1980, Vol. 3, pp. 179-186.

Teczan, S.S. and Ipek, M. (1973), "Long Distance Effects of the 28 March 1970 Gediz Turkey Earthquake," Earthquake Engineering and Structural Dynamics, Vol. 1, 1973, pp. 203-215.

Teczan, S.S., Seed, H.b., Whitman, R.V., Serff, N., Christian, J.T., Durgoglou, H.T. and Yegian, M.K. (1977), "Resonant Period Effects in the Gediz, Turkey Earthquake of 1970," Earthquake Engineering and Structural Dynamics," Vol. 5, 1977, pp. 157-179.

Tokimatsu, K. (1984), Personal Communication with S. Liao, 1984.

Tokimatsu, K. and Yoshimi, Y. (1983), Empirical Correlation of Soil Liquefaction Based on SPT N-value and Fines Content," Soils and Foundations, Vol. 23, No. 4, Dec. 1983.

†Tohno, I. and Yasuda, S. (1978), "Liquefaction of the Ground During the 1978 Miyagiken-Oki Earthquake," Soils and Foundations, Vol. 21, No. 3, Sept. 1981, pp. 18-34.

*Tsuchida, H. et al. (1979). "The Damage to Port Structures by the 1978 Miyagiken-Oki Earthquake," Technical Note, Port and Harbour Research Institute, Ministry of Transport, No. 325, 1979 (in Japanese).

†Tsuchida, H., Iai, S. and Hayashi, S. (1980), "Analysis of Liquefactions During the 1978 Off Miyagi Prefecture Earthquake," Proc. 7th World Conf. on Earthquake Engineering, Istanbul, Vol. 3, pp. 211-218.

W.C.C. (1976), Data from Woodward Clyde Consultants Files reported by Yegian, 1976.

Whitman, R.V. (1971), "Resistance of Soil to Liquefaction and Settlement," Soils and Foundations, Vol. 11, No. 4, Dec. 1971, pp. 59-68.

Whitman, R.V., Yegian, M.K., Christian, J.T. and Tezcan, S.S. (1974), "Ground Motion Amplification Studies, Bursa, Turkey," Seismic Design Decision Analysis Report No. 16, M.I.T. CER77-58, Sept. 1974.

Xie, J. (1979), "Empirical Criteria of Sands Liquefaction," The 1976 Tangshan China Earthquake - Papers presented at the 2nd U.S. National Conf. on Earthquake Engr., Stanford University, 1979, published by Earthquake Engineering Research Institute, Berkeley, CA., March, 1980.

Yamanouchi, T. et. al. (1970), "Disaster Features in 1968 Ebino Earthquake from the Viewpoint of Soil Engineering," Soils and Foundations, 1970, Vol. 10, No. 2, pp. 129-144.

Ye, Y. and Liu, X. (1979), "Experience in Engineering from Earthquake in Tangshan and Urban Control of Earthquake Disaster," The 1976 Tangshan Earthquake - Papers presented at the 2nd U.S. National Conf. on Earthquake Eng. Stanford Univ., 1979, published by Earthquake Engineering Research Institute, Berkeley, Calif. March 1980.

Yegian, M.K. (1976), "Risk Analysis for Earthquake-Induced Ground Failure by Liquefaction," Ph.D. Thesis, Dept. of Civil Eng., Mass. Inst. of Technology, Cambridge, Mass., May 1976.

Yegian, M.K. and Vitelli, B.M. (1981), "Probabilistic Analysis for Liquefaction," Report No. CE-81-1, Dept. of Civil Engineering Northeastern University, Boston, Mass., March 1981.

Yoshimi, Y. (1970), "An Outline of Damage During the Tokachioki Earthquake," Soils and Foundations, Vol. 10, No. 2, pp. 1-14.

Yoshimi, Y. and Akagi, T. (1968), News Report, "Tokachioki Earthquake of 1968", Soils and Foundations, Vol. 8, No. 3, Sept. 1968, pp. 87-95.

Youd, T.L. (1984), Personal Communication with S. Liao, 1984.

Youd, T.L. and Bennett, M.J. (1983), "Liquefaction Sites, Imperial Valley, California," Jour. of Geotech. Engr., ASCE, Vol. 109, No. 3, March 1983, pp. 440-457.

Youd, T.L. and Keefer, D.K. (1984), "Liquefaction During the 1977 San Juan, Argentina Earthquake: Site and Regional Studies," U.S. Geological Survey Publication (in press).

Youd, T.L. and Hoose, S.N. (1976), "Liquefaction During 1906 San Francisco Earthquake," Journal of the Geotechnical Engineering Division, ASCE, Vol. 102, No. GT5, May 1976, pp. 425-439.

Youd, T.L. and Hoose, S.N. (1978), "Historical Ground Failures in Northern California Associated with Earthquakes," U.S. Geological Survey Professional Paper 993, 1978.

Youd, T.L. and Wieczorek, G.F. (1984), "Liquefaction During the 1981 and Previous Earthquakes Near Westmoreland, California," Open File Report 84-680, U.S. Geological Survey, Menlo Park, Calif., 1984.

Zhou, S.G. (1981), "Influrence of Fines on Evaluating Liquefaction of Sand by CPT," Proc. Int'l Conf. on Recent Advances in Geotech. Earthquake Eng. and Soil Dynamics, St.Louis, Missouri, Vol. 1, pp.167-172.

---

* Indicates reference not available to the writer.

† Indicates supplmentary reference not used in catalog compilation, but may be of use in future studies.

APPENDIX B

COMPUTER PROGRAM LISTINGS

This appendix contains listings of the weighted binary logistic regression prcgram (WLOGIT) and the nonparametric estimation proqram (PGR) used in this study. Neither program is purported to be completely general, alouth WLOGIT has been used successfully by others. Other algorithms described in this study are extensions or modifications of these two programs. The logic of the programs and their proper usage are described by comment statements in the listings. The programs are coded in FORTRAN 77.

B.1 WLOGIT

WLOGIT is a subroutine, for which the user must write a simple main program to read the data and then call WLOGIT. The data is passed to WLOGIT through common blocks. Control options for WLOGIT are passed through the subroutine arguments. An example main program (LOMAIN) is also listed. Subroutine WLOGIT calls subroutine INVDET, which is a matrix inversion program obtained from Hornbeck (1975). Subroutine WLOGIT and INVDET are self-contained and do not rely on external software packages.

B.2 PGR

PGR is a program to calculate the probability $\hat{P}(Y=1)$ on a 2-dimensional grid of points and to plot the results in terms of contours of equal $\hat{P}$. This program was implemented on a Vax 11/782 computer using

the graphics subroutine package PENPLOT (Version 2). Availability of subroutines from the IMSL program library is also required. The program listing is provided here mainly as a guide to those who may wish to implement the nonparametric estimation procedure on other computer systems. The non-library subroutines of PGR are SMGR, KEREST, NEIBOR, SLTOUR, CONTUR, SLBOX, and DATPLT. Subroutines other than these are part of either PENPLOT or IMSL.

```
C     LOMAIN  --  MAIN PROGRM TO RUN WLOGIT
C     4/19/85
C
      COMMON/BLK1/NVAR,NCASE,Y,X,W
      COMMON/BLK2/VARID
      COMMON/OUT1/BETA,HESS
C
      REAL*4 Y(400),X(400,30),W(400)
      REAL*8 BETA(30),HESS(30,30)
      CHARACTER*12 VARID(30)
      INTEGER ICASE(400)
C
C     READ CONTROL AND VARIABLE ID  FILE
C     ----------------------------------
      READ (1,*) NVAR, NCHK, IPRNT  ! NCHK=No.OF CHECK OUTPUT CASES
                                    ! IPRNT=PRINT OPTION FOR WLOGIT
      READ (1,*) (VARID(I),I=1,NVAR)
C
C     READ THE EDITED DATA FILE
C     -------------------------
      READ(2,*) NCASE
C
      READ(2,*) (ICASE(I),Y(I),(X(I,J),J=1,NVAR),I=1,NCASE)
C
C     OUTPUT TO CHECK VARIABLES
C     -------------------------
      WRITE(3,50) ((ICASE(I),Y(I),(X(I,J),J=1,11)),I=1,NCHK)
   50 FORMAT(I8,12F8.3)
C
      CALL WLOGIT(0,IPRNT)
C
      STOP
      END
```

```
C     ================================
      SUBROUTINE WLOGIT(IOPT1,IPRNT)
C     ================================
C                                                    S.L. 6/18/1984
C                                                    S.L. 12/17/1984
C                                                    S.L. 1/12/1985
C                                                    S.L. 5/6/1985
C
C     SUBROUTINE VERSION OF PROGRAM FOR BINARY LOGIT REGRESSION
C     WITH WEIGHTED LOG-LIKELIHOOD FUNCTION
C
C     VARIABLES PASSED THROUGH ARGUMENTS
C       IOPT1     - WEIGHT OPTION:
C                   IOPT1 = 0  RESULTS IN UNWEIGHTED ANALYSIS
C                              OTHERWISE RESULTS IN WEIGTHED ANALYSIS
C
C       IPRNT     - PRINT OPTION:
C                   IPRNT = 0  ALL RESULTS PRINTED
C                         = 1  VARIANCE-COVARIANCE MATRIX NOT PRINTED
C                         = 2  ITERATION LOG *ALSO* NOT PRINTED
C                         = 3  PRINT HEADER AND MLRI ONLY
C                         = 9  NO RESULTS PRINTED
C
C     VARIABLES TO BE PASSED TO SLOGIT THROUGH COMMON BLOCK /BLK1/
C       NVAR      - NO. OF EXPLANATORY VARIABLES (UP TO 30)
C       NCASE     - NO. OF CASES (UP TO 300)
C       Y(I)      - RESPONSE VARIABLES FOR ITH CASE, I=1,NCASE
C       X(I,J)    - EXPLANATORY VARIABLES FOR ITH CASE,I=1,NCASE
C       W(I)      - LOG-LIKELIHOOD WEIGHTS FOR ITH CASE, I=1,NCASE
C
C     VARIABLES TO BE PASSED TO WLOGIT THROUGH COMMON BLOCK /BLK2/
C       VARID(I)  - NAMES OF EXPLANATORY VARIABLES, I=1,NVAR
C                   (CHARACTER*9)
C
C     GLOSSARY OF KEY VARIABLES IN PROGRAM:
C       BETA(*)   - VECTOR OF BETA COEFFICIENTS
C       CHI       - CHI-SQUARED STATISTIC
C       CONV1     - CONVERGENCE INDICATOR 1
C       CONV2     - CONVERGENCE INDICATOR 2
C       DETM      - DETERMINANT OF HESS
C       DOFCOR    - DEGREES-OF-FREEDOM CORRECTION FACTOR FOR VARIANCE-
C                   COVARIANCE MATRIX
C       DTNRM     - DETM DIVIDED BY EUCLIDIAN NORM OF HESS
C       DBETA(*)  - VECTOR OF INCREMENTAL VALUES OF BETA IN NEWTON-
C                   RAPHSON ITERATION
C       DLIKE(*)  - VECTOR OF DERIVATIVES OF LOG-LIKELIHOOD
C       HESS(*,*) - NEGATIVE OF THE HESSIAN MATRIX (2ND DERIVATIVES
C                   OF LOG-LIKELIHOOD FUNCTION)
C                 - ALSO USED TO STORE THE INVERSE OF HESS,WHICH
C                   AT CONVERGENCE IS THE VARIANCE-COVARIANCE
C                   MATRIX OF THE BETAS
C       ICNTRL    - VARIABLE FOR CONTROLLING FLOW OF PROGRAM
C       NCP       - NO. OF CASES CORRECTLY  PREDICTED BY LOGIT MODEL
C       NPAR      - NUMBER OF PARAMETERS IN A MODEL = NVAR + 1
C       NSTOP     - MAX NO. OF NEWTON-RAPHSON ITERATIONS
C       NYZ       - NO. OF CASES Y=0.
C       NY1       - NO. OF CASES Y=1.
C       PCP       - PERCENT CORECTLY PREDICTED STATISTIC
C       PNYZ      - PERCENT OF CASES Y=0.
C       PNY1      - PERCENT OF CASES Y=1.
```

```
C       TOL1       - CONVERGENCE CRITERION FOR CONV1
C       TOL2       - CONVERGENCE CRITERION FOR CONV2
C       P          - PROBABILITY OF Y=1.0 (FAILURE) FOR GIVEN CASE
C       STDERR(*)  - VECTOR OF STANDARD ERROR OF ESTIMATE FOR BETA PARAMETERS
C       T(*)       - VECTOR OF T-STATISTICS FOR BETA PARAMETERS
C       XLIKE      - LOG-LIKELIHOOD FUNCTION
C       XLIKEC     - LOG-LIKELIHOOD AT BETA=CONSTANT
C       XLIKEZ     - LOG-LIKELIHOOD AT BETA=0
C       XLRC       - LOG LIKELIHOOD RATIO STATISTIC AT BETA=CONST
C       XLRI       - LIKELIHOOD RATIO INDEX
C       XLRZ       - LOG-LIKELIHOOD RATIO STATISTIC AT BETA=0
C       XMLRI      - MODIFIED LIKELIHOOD RATIO INDEX
C       VARID(*)   - VECTOR OF NAMES OF EXPLANATORY VARIABLES
C
C                    ----------------------------
C
      COMMON/BLK1/NVAR,NCASE,Y,X,W
      COMMON/BLK2/VARID
      COMMON/OUT1/BETA,HESS
C
C DECLARATION OF ARRAYS REFERENCED THROUGH COMMON BLOCKS
      REAL*4 Y(400),X(400,30),W(400)
      REAL*8 BETA(30),HESS(30,30)
      CHARACTER*12 VARID(30)
C
C
C DECLARATION OF ARRAYS INTERNAL TO SUBROUTINE
      REAL*4 T(30),STDERR(30)
      REAL*8 ARG,P,XLIKE,DETM,DTNRM
      REAL*8 DBETA(30),DLIKE(30)
C
C
C PRELIMINARIES
C -------------
      NPAR = NVAR + 1
      NSTOP = 15
      TOL1 = 0.0005
      TOL2 = 0.001
C
C IF IOPT1 = 0, SET WEIGHTS W(I) = 1.0
      IF(IOPT1.EQ.0) THEN
        DO 3 I=1,NCASE
        W(I) = 1.0
    3   CONTINUE
      ELSE
      END IF
C
C SET LAST ELEMENT OF ALL X-VECTORS TO 1.
      DO 5 I=1,NCASE
      X(I,NPAR) = 1.
    5 CONTINUE
C
C SET LAST ELEMENT OF VARID EQUAL TO 'CONST
      VARID(NPAR) = 'CONST      '
C
C INITIALIZE BETA VECTOR USING TRIAL VALUES = 0.
      DO 8 I=1,NPAR
      BETA(I) = 0.
    8 CONTINUE
C
```

```
C
C CALCULATING SOME BASIC DATA
C
      NYZ = 0
      NY1 = 0
      DO 10  I=1,NCASE
      IY = NINT(Y(I))
      IF (IY.EQ.0) THEN
        NYZ = NYZ + 1
      ELSE
        NY1 = NY1 + 1
      END IF
   10 CONTINUE
      PNYZ = FLOAT(NYZ)*100./FLOAT(NCASE)
      PNY1 = FLOAT(NY1)*100./FLOAT(NCASE)
C
C
C PRINTING HEADER AND BASIC DATA FOR OUTPUT
C
      IF (IPRNT.GE.9) GOTO 2050
      IF (IPRNT.LE.2) WRITE(6,2011)
      WRITE(6,11)
      WRITE(6,12)
      WRITE(6,13) NYZ,PNYZ
      WRITE(6,14) NY1,PNY1
      WRITE(6,15) NCASE
      WRITE(6,16) NVAR,(VARID(I),I=1,NVAR)
      IF (IPRNT.LE.1) WRITE(6,17)
C
 2011 FORMAT('1')
   11 FORMAT(    ///,4X,'************************************'
     1             / 4X,'* WEIGHTED BINARY LOGIT ANALYSIS *'
     2             / 4X,'************************************'//)
   12 FORMAT(5X,'   ALTERNATIVE      NO. OF OCCURENCES    % OCCURENCE'
     1      /5X,'-----------------   -----------------    -----------')
   13 FORMAT(5X,'Y=0 (NO FAILURE)',8X,I4,15X,F6.2)
   14 FORMAT(5X,'Y=1 (FAILURE)    ',8X,I4,15X,F6.2)
   15 FORMAT(//5X,          'TOTAL NO. OF CASES   =  ',I4)
   16 FORMAT(//5X,I2,X,     'EXPLANATORY VARIABLE(S):',4(2X,A9),
     1                                  (/32X,4(2X,A9)))
   17 FORMAT(///5X,'ITERATION LOG'
     1         /5X,'-------------'
     2         /5X,'IT#  LOG LIKELIHOOD  CHI-SQUARE    RMS DBETA
     3 CONV1       CONV2 - %        DETM        DTNRM'
     4         /5X,'---  --------------  ----------   -----------    -
     5--------    ----------      --------     -------')
C
 2050 CONTINUE
C
C
C START OF ITERATIONS TO MAXIMIZE XLIKE
C -------------------------------------
C
C INITIALIZE ITERATION COUNTER AND PROGRAM FLOW CONTROL VARIABLE
C
      ITER = 0
      ICNTRL = 0
C
C INITIALIZE COUNTER FOR PCP STATISTIC
C
```

```
      NCP = 0
C
C
C BRANCH  OF OVERALL LOOP OF NEWTON-RAPHSON ITERATIONS
C
  100 CONTINUE
C
C INITIALIZE  OF XLIKE, DLIKE,DBETA, AND HESS FOR EACH ITERATION
C
      XLIKE = 0.
C
      DO 101 I=1,NPAR
      DLIKE(I) = 0.
      DBETA(I) = 0.
  101 CONTINUE
C
      DO 120 I=1,NPAR
      DO 110 J=1,NPAR
      HESS(I,J) = 0.
  110 CONTINUE
  120 CONTINUE
C
C
C INITIALIZE CHI STATISTIC FOR EACH ITERATION
C
      CHI = 0.
C
C
C FORMATION OF XLIKE, DLIKE, AND HESS SUMS
C ----------------------------------------
C
C SUMMATION LOOP I OVER ALL CASES
C
  200 DO 500 I=1,NCASE
C
C CALCUALTE P FOR THE ITH CASE
C   NOTE: ARG = BETA(1)*X(I,1) + ... + BETA(NPAR)*X(I,NPAR)
C    AND: BETA(NPAR) IS THE CONSTANT COEFFICIENT, X(I,NPAR)=1.0
C
      ARG =0.
      DO 210 J=1,NPAR
      ARG = ARG + BETA(J)*X(I,J)
  210 CONTINUE
C
C     INTERVENTION TO PREVENT OVERFLOW/UNDERFLOW
C      * NOTE EXP(87.5) = 1.0E38
C
      IF (ARG.GT. 87.5) ARG =  87.5
      IF (ARG.LT.-87.5) ARG = -87.5
C
      P = 1.0/(1.0 + EXP(-ARG))
C
C
C SUM FOR XLIKE
C
      IY = NINT(Y(I))
      IF(IY.EQ.1) THEN
        XLIKE = XLIKE + W(I)*LOG(P)
      ELSE
        XLIKE = XLIKE + W(I)*LOG(1.0 - P)
```

```
      END IF
C
C SAVING THE INITIAL VALUE OF XLIKE FOR TEST STATISTICS
C
      IF (ITER.EQ.0) THEN
        XLIKEZ = XLIKE
      ELSE
      END IF
C
C SUM FOR DLIKE
C
      DO 310 J=1,NPAR
      DLIKE(J) = DLIKE(J) + W(I)*(Y(I) - P)*X(I,J)
  310 CONTINUE
C
C SUM FOR HESS
C
      DO 420 J=1,NPAR
      DO 410 K=1,NPAR
      HESS(J,K)= HESS(J,K) + W(I)*X(I,J)*X(I,K)*P*(1.0 - P)
  410 CONTINUE
  420 CONTINUE
C
C SUM FOR CHI STATISTIC
C
C     BUT FIRST - INTERVENTION TO PREVENT DIVISION BY ZERO
C
      IF (P.LT.0.0000001) P = 0.0000001
      IF (P.GT.0.9999999) P = 0.9999999
C
C
      CHI1 = ((Y(I) - P)**2)/P
      CHI2 = (((1.0 - Y(I)) - (1.0 - P))**2)/(1.0 - P)
      CHI = CHI + CHI1 + CHI2
C
C
C DECISION BRANCH FOR CONTROLLING PROGRAM FLOW (EFFECTIVE ONLY
C AFTER LAST ITERATION OF NEWTON-RAPHSON ALGORITHM)
C
C   ICNTRL=0 CAUSES "NORMAL" FLOW OF PROGRAM
C   ICNTRL=1 CAUSES THE PERCENT CORRECTLY PREDICTED STATISTIC
C            TO BE CALCULATED USING THE FINAL BETA VALUES
C
      IF (ICNTRL.EQ.0) GOTO 500
C
C ESTIMATE PERCENT CORRECTLY PREDICTED
C
      IF(P-0.5) 451,452,452
  451 IF (IY.EQ.0) GOTO 453
      GOTO 460
  452 IF (IY.EQ.1) GOTO 453
      GOTO 460
  453 NCP = NCP + 1
C
  460 CONTINUE
C
C
C END OF LOOP OVER CASES --  FROM STMT 200
C
  500 CONTINUE
```

```
C
C
C APPLICATION OF NEWTON-RAPHSON FORMULA
C -------------------------------------
C
  INVERTING HESS
C
      CALL INVDET(HESS,NPAR,DTNRM,DETM)
C
C DECISION BRANCH FOR CONTROLLING FLOW OF PROGRAM (EFFECTIVE
C ONLY AFTER LAST ITERATION OF NEWTON-RAPHSON ALGORITHM)
C
C    ICNTRL=0 CAUSES "NORMAL" FLOW OF PROGRAM
C    ICNTRL=1 CAUSES BRANCH TO CALCULATIONS OF TEST STATISTICS,
C             AVOIDING CALCULATIONS OF NEW BETA VECTOR
C
   501 IF (ICNTRL.EQ.1) GOTO 900

C MULTIPLY HESS WITH DLIKE TO GET DBETA
      DO 620 I=1,NPAR
      DO 610 J=1,NPAR
      DBETA(I) = DBETA(I) + HESS(I,J)*DLIKE(J)
   610 CONTINUE
   620 CONTINUE
C
C
C INCREMENT BETA VALUES BY DBETA
      DO 640 I=1,NPAR
   640 BETA(I) = BETA(I) + DBETA(I)
C
C ITERATION COMPLETED, INCREMENT ITER
      ITER = ITER + 1
C
C
C CONVERGENCE CRITERION CALCULATIONS
C ----------------------------------
C
C COMPUTE CONV1 = [DLIKE]'[HESS-INVERSE][DLIKE]
C   NOTE: [HESS-INVERSE][DLIKE] PREVIOUSLY CALCULATED
C            AND STORED IN DBETA VECTOR
C
      CONV1 =0.
      DO 710 I=1,NPAR
      CONV1 = CONV1 + DLIKE(I)*DBETA(I)
   710 CONTINUE
C
C
C COMPUTE  CONV2 = MAX PERCENT CHANGE IN BETA
C
      CONV2 = 0.
      DO 720 I=1,NPAR
      RBETA = ABS(DBETA(I)/BETA(I))*100.0
      CONV2 = MAX(CONV2,RBETA)
   720 CONTINUE
C
C
C CALCULATE RMS OF DBETA VECTOR (RMSDB)
C NOTE: THIS IS NOT USED AS A CONVERGENCE CRITERION TO
C       TERMINATE PROGRAM, BUT IS CALCULATED AND PRINTED
C       AS AN AUXILIARY INDICATOR OF CONVERGENCE
```

```
C
      RMSDB = 0.
      DO 730 I=1,NPAR
      RMSDB = RMSDB + DBETA(I)**2
  730 CONTINUE
      RMSDB = SQRT(RMSDB)
C
C
C
C ITERATION LOG PRINTOUT
C ----------------------
C
      IF (IPRNT.LE.1) THEN
       WRITE (6,801) ITER,XLIKE,CHI,RMSDB,CONV1,CONV2,DETM,DTNRM
      ELSE
      END IF
C
  801 FORMAT(/5X,I2,2X,7(2X,G12.5)/)
C
C
C TERMINATING NEWTON-RAPHSON ITERATIONS
C -------------------------------------
C
      IF (ITER.GE.NSTOP) GOTO 899
      IF (CONV1.GT.TOL1) GOTO 100
      IF (CONV2.GT.TOL2) GOTO 100
C
C
  899 CONTINUE
      ICNTRL = 1
      GOTO 100
C
C BRANCH FROM STMT 501
C
  900 CONTINUE
C
C
C CALCULATION OF TEST STATISTICS
C ------------------------------
C
      XNCASE = FLOAT(NCASE)
      XNPAR = FLOAT(NPAR)
      XNCP = FLOAT(NCP)
      XNYZ = FLOAT(NYZ)
      XNY1 = FLOAT(NY1)
C
C PERCENT CORRECTLY PREDICTED
C
      PCP = (XNCP/XNCASE)*100.
C
C
C LOG-LIKELIHOOD WITH CONSTANT COFFICIENT ONLY - XLIKEC
C
      XLIKEC = 0.
      XLOGP1 = LOG(PNY1/100.0)
      XLOGPZ = LOG(PNYZ/100.0)
C
      DO 920 I=1,NCASE
      IY = NINT(Y(I))
      IF (IY.EQ.1) THEN
```

```
        XLIKEC = XLIKEC + W(I)*XLOGP1
      ELSE
        XLIKEC = XLIKEC + W(I)*XLOGPZ
      END IF
  920 CONTINUE
C
C
C LOG-LIKELIHOOD RATIOS LRZ AND LRC
C
      XLRZ = 2.0*(XLIKE - XLIKEZ)
      XLRC = 2.0*(XLIKE - XLIKEC)
C
C
C LIKELIHOOD RATIO INDICES
C
      XLRI = 1.0 - (XLIKE/XLIKEZ)
      XMLRI = 1.0 - (XLIKE-XNPAR/2.0)/XLIKEZ
C
C
C T-STATISTIC FOR BETA PARAMETERS
C    NOTE: THE DIAGONAL ELEMENTS OF -[HESS-INVERSE] AT
C          CONVERGENCE ARE THE VARIANCES OF THE BETAS
C
C CALCULATE THE ADJUSTED VARIANCE-COVARIANCE MATRIX, I.E.,
C    CORRECT FOR DEGREES OF FREEDOM
C
      DOFCOR = FLOAT(NCASE)/FLOAT(NCASE - NPAR)
C
      DO 960 I=1,NPAR
      DO 950 J=1,NPAR
      HESS(I,J) = DOFCOR*HESS(I,J)
  950 CONTINUE
      STDERR(I) = SQRT(HESS(I,I))
      T(I) = BETA(I)/STDERR(I)
  960 CONTINUE
C
C
C
C FINAL OUTPUT
C ------------
C
      IF (IPRNT.GE.3) GOTO 3005
      IF (IPRNT.LE.1) WRITE(6,1008)
      WRITE(6,1010) IOPT1
      IF (ITER.GE.NSTOP) WRITE (6,1234)
      WRITE(6,1020) ((VARID(I),BETA(I),STDERR(I),T(I)),I=1,NPAR)
      WRITE(6,1030) XLIKEZ
      WRITE(6,1040) XLIKEC
      WRITE(6,1050) XLIKE
      WRITE(6,1060) XLRZ
      WRITE(6,1070) XLRC
      WRITE(6,1080) XLRI
      WRITE(6,1090) XMLRI
      WRITE(6,1100) CHI
      WRITE(6,1110) PCP
      IF (IPRNT.LE.0) THEN
       WRITE(6,1111)
       DO 1001 I=1,NPAR
       WRITE(6,1120) (HESS(I,J),J=1,NPAR)
 1001  CONTINUE
```

```
       WRITE(6,1130) DETM,DTNRM
      ELSE
      END IF
      GOTO 9999
C
 3005 CONTINUE
      WRITE(6,1140)
      IF (ITER.GE.NSTOP) WRITE (6,1234)
      WRITE(6,1090) XMLRI
C
C
 1008 FORMAT('1')
 1010 FORMAT(   ///5X,'COEFFICIENT SUMMARY - WEIGTED LOG-LIKELIHOODS',
     1 2X,'IOPT1 = ',I1,/
     2              5X '-------------------------------------------',
     3 2X,'---------'/
     4  //5X,'  VARIABLE          BETA        STAND. ERR    T-STATISTIC'
     5   /5X,'------------   ------------  ------------  -----------')
 1020 FORMAT((5X,A12,3X,3(G12.5,2X)/))
 1030 FORMAT(//5X,'LOG-LIKELIHOOD (BETA = 0)          L(0) = ',G12.5)
 1040 FORMAT(5X,  'LOG-LIKELIHOOD (BETA = CONST)      L(C) = ',G12.5)
 1050 FORMAT(5X,  'LOG-LIKELIHOOD AT CONVERGENCE      L(B) = ',G12.5)
 1060 FORMAT(//5X,'LIKELIHOOD RATIOS LRZ = -2(L(0)-L(B)) = ',G12.5)
 1070 FORMAT(5X,  '                  LRC = -2(L(C)-L(B)) = ',G12.5)
 1080 FORMAT(//5X,'LIKELIHOOD RATIO INDEX RHO**2 = XLRI  = ',G12.5)
 1090 FORMAT(5X,  'MODIFIED LRI       RHO-BAR**2 = XMLRI = ',G12.5)
 1100 FORMAT(5X,  'CHI-SQUARED STATISTIC             CHI  = ',G12.5)
 1110 FORMAT(5X,  'PERCENT CORRECTLY PREDICTED       PCP  = ',F6.2)
 1111 FORMAT(//5X,'DOF-ADJUSTED HESS-INVERSE (VARIANCE-COVARIANCE)
     1  MATRIX'/)
 1120 FORMAT(5X,6(2X,G12.5))
 1130 FORMAT(/5X,'DETM = ',G15.5,5X,'DTNRM = ',G12.5)
 1140 FORMAT(//)
 1234 FORMAT(/5X,'          ********  NO CONVERGENCE  ********'/)
C
C
 9999 RETURN
      END
```

```
C INVDET   ------  version for max. 30 by 30 matrix
C SUBROUTINE TO INVERT A MATRIX
C REF: HORNBECK,R.W.:"NUMERICAL METHODS", QUANTUM PUBLISHERS,1975
C NOTE THE CALCULATION OF EUCLIDEAN NORM (PD) IN HORNBECK'S
C PROGRAM IS INCORRECT AND HAS BEEN APPROPRIATELY MODIFIED
C PROGRAM IS NOW CHANGED TO DOUBLE PRECISION
C ALSO, ALL RELEVANT VARIABLES HAVE BEEN CHANGED TO DOUBLE PRECISION
C
      SUBROUTINE INVDET(C,N,DTNRM,DETM)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION C(30,30),J(60)
      PD=0.0
      DO 124 L=1,N
      DD=0.0
      DO 123 K=1,N
  123 DD=DD+C(L,K)*C(L,K)
      PD=PD+DD
  124 CONTINUE
      PD=DSQRT(PD)
      DETM=1.0
      DO 125 L=1,N
  125 J(L+20)=L
      DO 144 L=1,N
      CC=0.0
      M=L
      DO 135 K=L,N
      IF((DABS(CC)-DABS(C(L,K))).GE.0.0) GO TO 135
  126 M=K
      CC=C(L,K)
  135 CONTINUE
  127 IF(L.EQ.M) GO TO 138
  128 K=J(M+20)
      J(M+20)=J(L+20)
      J(L+20)=K
      DO 137 K=1,N
      S=C(K,L)
      C(K,L)=C(K,M)
  137 C(K,M)=S
  138 C(L,L)=1.0
      DETM=DETM*CC
      DO 139 M=1,N
  139 C(L,M)=C(L,M)/CC
      DO 142 M=1,N
      IF(L.EQ.M) GO TO 142
  129 CC=C(M,L)
      IF(CC.EQ.0.0) GO TO 142
  130 C(M,L)=0.0
      DO 141 K=1,N
  141 C(M,K)=C(M,K)-CC*C(L,K)
  142 CONTINUE
  144 CONTINUE
      DO 143 L=1,N
      IF(J(L+20).EQ.L) GO TO 143
  131 M=L
  132 M=M+1
      IF(J(M+20).EQ.L) GO TO 133
  136 IF (N.GT.M) GO TO 132
  133 J(M+20)=J(L+20)
      DO 163 K=1,N
      CC=C(L,K)
      C(L,K)=C(M,K)
```

```
163 C(M,K)=CC
    J(L+20)=L
143 CONTINUE
    DETM=DABS(DETM)
    DTNRM=DETM/PD
    RETURN
    END
```

```
C     PGR - KERNEL BINARY REGRESSION -- PROBABILITY ESTIMATES ON A GRID
C     MAIN PROGRAM TO READ DATA, RUN SMPR, AND PLOT RESULTS
C
C *** INPUT GLOSSARY:
C
C --- PLOT LABELS
C     PLAB   =  PLOT LABEL
C     XLAB   =  X-AXIS LABEL OF PLOT
C     YLAB   =  Y-AXIS LABEL OF PLOTS
C
C --- KERNEL SMOOTHING PARAMETERS
C     ISIZE  =  SMOOTHER SIZE OPTION INDICATOR
C                  0 = USE CONSTANT SIZE KERNEL
C                  1 = USE VARIABLE SIZE KERNEL TO BE CALCUALTED AS
C                      A LINEAR FUNCTION OF THE DISTANCE TO THE KTH
C                      NEAREST NEIGHBOR
C     INTK   =  OPTION CODE FOR INTEGRATED SMOOTHER
C                  0 = USE STANDARD KERNEL SMOOTHER
C                  1 = USE INTEGRATED KERNEL SMOOTHER
C     KTH    =  SPECIFIED KTH NEAREST NEIGHBOR USED AS A DISTANCE
C               MEASURE (DOESN'T MATTER WHAT IS SPECIFIED IF ISIZE=0)
C     CSIZE  =  SIZE CONSTANT OF VARIABLE KERNEL SMOOTHER
C               (DOESN'T MATTER WHAT IS SPECIFIED IF INTK=0)
C     SIGX   =  STD DEV ALONG X-AXIS OF KERNEL(DOESN'T MATTER WHAT
C               IS SPECIFIED IF ISIZE=1)
C     RYX    =  RATIO OF 1 UNIT OF Y-AXIS BANDWIDTH TO 1 UNIT OF X-AXIS
C               BANDWIDTH IN THE SMOOTHING SPACE (NATURAL, LOG-NATURAL,
C               OR LOG-LOG), NOT JUST THE NATURAL SCALE
C     RHO    =  CORRELATION COEFFICIENT OF KERNEL SMOOTHER
C     BEXP   =  POWER COEFFICIENT FOR DISTANCE TO KTH NEAREST NEIGHBOR
C               (DOESN'T MATTER WHAT IS SPECIFIED IF ISIZE=0)
C
C --- PLOTTING/OUTPUT AND CONTROL OPTIONS
C
C     ISYM   =  OPTION CODE FOR PLOTTING SYMBOLS
C                  0 = DISTINGUISH BETWEEN RESP=1 AND RESP=0
C                  1 = PLOT SAME SYMBOL (+) FOR BOTH
C     VINT   =  CONTOUR INTERVAL USED IN PLOTS
C     ILOGX  =  OPTION CODE FOR TRANSFORMING X-COORDINATES TO LOG SCALE
C               FOR  SMOOTHING
C                  0 = USE NATURAL SCALE
C                  1 = USE LOG10 TRANSFORMED SCALE
C     ILOGY  =  SAME CODE AS ILOGX FOR Y-COORDINATES
C     IPLOT  =  OPTION CODE FOR PLOTS GENERATED
C                  0 = PLOT ONLY NATURAL SCALE PLOT  ** ASSUMES NO
C                      LOG SCALE TRANSFORMS USED IN SMOOTHING **
C                  1 = PLOT ONLY LOG SCALE PLOT
C                  2 = PLOT LOG SCALE PLOT FIRST,AND THEN NATURAL SCALE
C     IWRT   =  OPTION CODE FOR WRITING RESULTS (P-VALUES) IN A FILE (UNIT 6)
C               TO BE USED FOR FURTHER CALCULATIONS OR RE-PLOTTING
C               AT SOME LATER TIME
C                  0 = DONT WRITE FILE
C                  1 = WRITE FILE
C
C --- PLOT SPECS FOR NATURAL AND LOG10 SCALE PLOT
C      LOGX,XLO,XHI,NXDIV ETC.--- SEE SUBROUTINE SLBOX FOR DETAILS
C      XL,XR,YB,YT, ETC. = LEFT,RIGHT,BOTTOM,TOP CLIPPING LIMITS FOR PLOTS
C
C --- GRID SPECS IN UNITS OF NATURAL SCALE
C     IMAX   =  MAX NO. OF GRID COORDINATES IN X-DIRECTION
```

```
C     JMAX  =  MAX NO. OF GRID COORDINATES IN Y-DIRECTION
C     XG(I)  =  X GRID COORDINATES, I=1,IMAX
C     YG(J)  =  Y GRID COORDINATES, J=1,JMAX
C
C --- FORMAT STMT FOR OUTPUT
C     FMT8   = FORMAT STMT FOR UNIT 8  -- THE DATA CHECK OUTPUT
C
C --- DATA INPUT FROM UNIT 5
C     NDATA  =  NUMBER OF DATA TO BE READ AND USED IN SMOOTHING
C     RESP(K)=  RESPONSE (0/1) VARIABLE 1=FAILURE 0=NONFAILURE
C     XD(K)  =  X-COORDINATE OF DATA INPUT, K=1,NDATA
C     YD(K)  =  Y-COORDINATE OF DATA INPUT, K=1,NDATA
C
C
C *** GLOSSARY OF SOME OTHER VARIABLES USED IN THE PROGRAM
C
C     DK(I,J)  =  DISTANCE TO KTH NEAREST NEIGHBOR OF GRID POINT (I,J)
C     P(I,J)   =  VALUE OF ESTIMATE OF FUNCTION AT (I,J)
C     SX(I,J)  =  STD. DEV. OF X FOR KERNEL AT GRID POINT (I,J)
C     SY(I,J)  =  STD. DEV. OF Y FOR KERNEL AT GRID POINT(I,J)
C     SKIP     =  A DUMMY CHARACTER*80 VARIABLE, SO HEADERS CAN BE USED
C                 IN THE DATA FILE
C
      COMMON/SPEC/IMAX,JMAX,NDATA
      COMMON/BASE/RESP,XD,YD
      COMMON/GRID/XG,YG,P
      COMMON/SXYR/SX,SY,R,DK
C
      DIMENSION RESP(400), XD(400), YD(400)
      DIMENSION XG(50), YG(50), P(50,50)
      DIMENSION SX(50,50), SY(50,50), R(50,50), DK(50,50)
C
      CHARACTER*60 PLAB,XLAB,YLAB
      CHARACTER*80 FMT8
      CHARACTER*80 SKIP
C
C     READ CONTROL FILE
C     -----------------
      READ(1,*) SKIP
C
C ----PLOT LABELS
      READ(1,*) SKIP
      READ(1,*) PLAB
      READ(1,*) XLAB
      READ(1,*) YLAB
C
C ----KERNEL SMOOTHING PARAMETERS
      READ(1,*) SKIP
      READ(1,*) ISIZE, INTK
      READ(1,*) SIGX, RYX, RHO
      READ(1,*) KTH, CSIZE, BEXP
C
C ----PLOTTING/OUTPUT AND CONTROL OPTIONS
      READ(1,*) SKIP
      READ(1,*) ISYM
      READ(1,*) VINT               ! CONTOUR INTERVAL DESIRED
      READ(1,*) ILOGX,ILOGY,IPLOT  ! X AND Y LOG-SCALING AND PLOT OPTIONS
      READ(1,*) IWRT               ! OPTION CODE TO WRITE OUTPUT FILE
C
C ----PLOT SPECS FOR NATURAL SCALE PLOT
```

```
      READ(1,*) SKIP
      READ(1,*) LOGX,XLO,XHI,NXDIV
      READ(1,*) LOGY,YLO,YHI,NYDIV
      READ(1,*) IOPTC,UL,UR,VB,VT
      READ(1,*) XL,XR,YB,YT              ! CLIPPING LIMITS
C
C ----PLOT SPECS FOR LOG SCALE PLOT
      READ(1,*) SKIP
      READ(1,*) LOGX1,XLO1,XHI1,NXDIV1
      READ(1,*) LOGY1,YLO1,YHI1,NYDIV1
      READ(1,*) IOPTC1,UL1,UR1,VB1,VT1
      READ(1,*) XL1,XR1,YB1,YT1          ! CLIPPING LIMITS
C
C ----GRID SPECS IN UNITS OF NATURAL SCALE
      READ(1,*) SKIP
      READ(1,*) IMAX,JMAX
      READ(1,*) (XG(I),I=1,IMAX)
      READ(1,*) (YG(J),J=1,JMAX)
C
C ----FORMAT STMTS FOR INPUT AND OUTPUT
      READ(1,*) SKIP
      READ(1,*) FMT8
C
C
C     READ THE DATA FROM UNIT 5
C     -------------------------
      READ(5,*) NDATA
C
      DO K=1,NDATA
      READ(5,*) RESP(K),XD(K),YD(K)
      END DO
C
C
C     SET CONSTANTS TO ZERO WHEN NOT APPLICABLE TO ESTIMATE
C     -----------------------------------------------------
      IF(ISIZE.EQ.0) THEN
       KTH = 0
       CSIZE = 0.
       BEXP = 0.
      ELSE
      END IF
C
C     INPUT CHECK
C     -----------
      WRITE(8,201) ILOGX, ILOGY, IPLOT, IWRT
      WRITE(8,202) ISIZE,INTK,SIGX,RYX,RHO,KTH,CSIZE,BEXP
      WRITE(8,203) (XG(I),I=1,IMAX)
      WRITE(8,204) (YG(J),J=1,JMAX)
      WRITE(8,205) (XD(K),YD(K),K=1,10)
      WRITE(8,206) VINT
  201 FORMAT (//' ILOGX = ',I1,'  ILOGY = ',I1, '  IPLOT = ',I1,
     1          '  IWRT = ',I1)
  202 FORMAT ( /' ISIZE = ',I1,'  INTK  = ',I1//
     1        '  SIGX = ',G11.5,'  RYX = ',G11.5,'  RHO = ',G11.5,//
     2        '  KTH =',I3,' CSIZE = ',G11.5,'  BEXP= ',G11.5)
  203 FORMAT ( /' X GRID COORDINATES:'/(8(X,G8.2)))
  204 FORMAT ( /' Y GRID COORDINATES:'/(8(X,G8.2)))
  205 FORMAT ( /' THE FIRST 10 DATA POINTS (X,Y) ARE:'/(8(X,G8.2)))
  206 FORMAT (//' CONTOUR INTERVAL VINT = ', G8.3/)
C
```

```
C
C     WRITING THE FIRST PART OF THE OUTPUT FILE IF DESIRED
C     ----------------------------------------------------
      IF (IWRT.EQ.1) THEN
       WRITE(6,*) IMAX,JMAX
       WRITE(6,*) (XG(I),I=1,IMAX)
       WRITE(6,*) (YG(J),J=1,JMAX)
      ELSE
      END IF
C
C     PLOT THE NATURAL SCALE VERSION OF THE DATA IF WORKING ONLY
C     IN THE NATURAL SCALE (THIS IS DONE HERE AS A DATA CHECK)
C     ----------------------------------------------------------
      IF (IPLOT.GT.0) GOTO 10
      CALL SLBOX(LOGX,XLO,XHI,NXDIV, LOGY,YLO,YHI,NYDIV,
     1             XLAB,YLAB,PLAB, IOPTC,MARGIN(UL,UR,VB,VT))
C
      CALL CLIP(XL,XR,YB,YT)
      CALL DATPLT(ISYM)
      GOTO 20
C
   10 CONTINUE
C
C     CONVERT THE DATA AND GRID COORDINATES TO TO LOG10 UNITS
C     -------------------------------------------------------
      IF (ILOGX.EQ.1) THEN
C
       DO K=1,NDATA
       XD(K) = LOG10(XD(K))
       END DO
C
       DO I=1,IMAX
       XG(I) = LOG10(XG(I))
       END DO
C
      ELSE
      END IF
C
      IF (ILOGY.EQ.1) THEN
C
       DO K=1,NDATA
       YD(K) = LOG10(YD(K))
       END DO
C
       DO J=1,JMAX
       YG(J) = LOG10(YG(J))
       END DO
C
      ELSE
      END IF
C
C     PLOT THE DATA IN THE LOG SCALE (DATA CHECK)
C     -------------------------------------------
      CALL SLBOX(LOGX1,XLO1,XHI1,NXDIV1, LOGY1,YLO1.YHI1,NYDIV1,
     1             XLAB,YLAB,PLAB, IOPTC1,MARGIN(UL1,UR1,VB1,VT1))
      CALL CLIP(XL1,XR1,YB1,YT1)
      CALL DATPLT(ISYM)
C
   20 CONTINUE
C
```

```
C     OBTAIN KERNEL-SMOOTHED ESTIMATES OF P
C     -------------------------------------
      CALL SMGR(ISIZE,INTK,SIGX,RYX,RHO,KTH,CSIZE,BEXP)
C
C
C     WRITE A CHECK FILE TO INTERPRET CONTOUR DATA
C     --------------------------------------------
      WRITE(8,40)
   40 FORMAT('1')
      WRITE(8,FMT8)  ((P(I,J),I=1,IMAX),J=JMAX,1,-1)
C
C     PLOT CONTOURS
C     -------------
      CALL SLTOUR(VINT)
C
      CALL ENDPLT
C
C
C     IF KERNEL SMOOTHED ESTIMATE WAS DONE IN THE LOG SCALE, NOW
C     PLOT THE DATA AND CONTOURS IN A NATURAL SCALE IF DESIRED
C     ========================================================
      IF (IPLOT.NE.2) GOTO 99     ! DON'T PLOT

C     CONVERT THE DATA AND GRID COORDINATES BACK TO NATURAL UNITS
C     ----------------------------------------------------------
      IF (ILOGX.EQ.1) THEN
C
       DO K=1,NDATA
       XD(K) = 10.0**XD(K)
       END DO
C
       DO I=1,IMAX
       XG(I) = 10.0**XG(I)
       END DO
C
      ELSE
      END IF
C
      IF (ILOGY.EQ.1) THEN
C
       DO K=1,NDATA
       YD(K) = 10.0**YD(K)
       END DO
C
       DO J=1,JMAX
       YG(J) = 10.0**YG(J)
       END DO
C
      ELSE
      END IF
C
C     PLOT DATA
C     ---------
      CALL LINE(0)
      CALL SLBOX(LOGX,XLO,XHI,NXDIV, LOGY,YLO,YHI,NYDIV,
     1             XLAB,YLAB,PLAB,  IOPTC,MARGIN(UL,UR,VB,VT))
C
      CALL CLIP(XL,XR,YB,YT)
      CALL DATPLT(ISYM)
C
```

```
C     PLOT CONTOURS
C     -------------
      CALL SLTOUR(VINT)
C
      CALL ENDPLT
C
   99 CONTINUE
C
      IF(IWRT.EQ.1) THEN
       DO I=1,IMAX
       WRITE(6,*) (P(I,J),J=1,JMAX)
       END DO
      ELSE
      END IF
C
      STOP
      END
```

```
C     ================================================================
      SUBROUTINE SMGR(ISIZE,INTK,SIGX,RYX,RHO,KTH,CSIZE,BEXP)
C     ================================================================
C
C     VARIABLE-SIZED SMOOTHER USING A BIVARIATE NORMAL KERNEL
C     SMOOTHER USED TO CALCULATE ESTIMATE OF P AT SPECIFIED
C     GRID POINTS
C
C     INPUT ARGUMENT GLOSSARY
C     -----------------------
C     ISIZE  =  SMOOTHER SIZE OPTION INDICATOR
C                  0 = USE CONSTANT SIZE KERNEL
C                  1 = USE VARIABLE SIZE KERNEL TO BE CALCUALTED AS
C                      A LINEAR FUNCTION OF THE DISTANCE TO THE KTH
C                      NEAREST NEIGHBOR (IN 2-DIMESIONS)
C     INTK   =  OPTION CODE FOR INTEGRATED KERNEL
C                  0 = USE STANDARD KERNEL
C                  1 = USE INTEGRATED KERNEL
C     SIGX   =  STD DEV BANDWIDTH ALONG X-AXIS OF KERNEL
C               *** FOR A VARIABLE KERNEL SIGX DEFINES THE X-SCALE OR
C                   "METRIC" FOR MEASURING THE DISTANCES TO THE KTH
C                   NEAREST NEIGHBOR ******
C     RYX    =  RATIO OF Y-AXIS BANDWIDTH TO X-AXIS BANDWIDTH
C     RHO    =  CORRELATION COEFFICIENT OF KERNEL SMOOTHER
C     KTH    =  SPECIFIED KTH NEAREST NEIGHBOR USED AS A DISTANCE
C               MEASURE
C     CSIZE  =  SIZE CONSTANT FOR VARIABLE KERNEL SMOOTHER
C     BEXP   =  POWER COEFFICIENT FOR DISTANCE TO KTH NEAREST NEIGHBOR
C
C     DK(I,J)  =  DISTANCE TO KTH NEAREST NEIGHBOR OF GRID POINT (I,J)
C
C     GLOSSARY OF SOME VARIABLES USED IN THE PROGRAM
C     ----------------------------------------------
C     IMAX  =  MAX NO. OF GRID COORDINATES IN X-DIRECTION
C     JMAX  =  MAX NO. OF GRID COORDINATES IN Y-DIRECTION
C     NDATA  =  NUMBER OF DATA POINTS
C
C     SX(I,J)  =  STD. DEV. BANDWIDTH OF X FOR KERNEL AT GRID POINT (I,J)
C     SY(I,J)  =  STD. DEV. BANDWIDTH OF Y FOR KERNEL AT GRID POINT (I,J)
C
C     XD(K)  =  X-COORDINATE OF DATA INPUT, K=1,NDATA
C     YD(K)  =  Y-COORDINATE OF DATA INPUT, K=1,NDATA
C     RESP(K)  =  RESPONSE (0/1) VARIABLE  1=FAILURE  0=NONFAILURE
C
C     XG(I)  =  X GRID COORDINATES, I=1,IMAX
C     YG(J)  =  Y GRID COORDINATES, J=1,JMAX
C     P(I,J)  =  VALUE OF ESTIMATE OF P AT (I,J)
C
C
      COMMON/SPEC/IMAX,JMAX,NDATA
      COMMON/BASE/RESP,XD,YD
      COMMON/GRID/XG,YG,P
      COMMON/SXYR/SX,SY
C
      DIMENSION RESP(400), XD(400), YD(400)
      DIMENSION XG(50), YG(50), P(50,50)
      DIMENSION SX(50,50), SY(50,50), DK(50,50)
      DIMENSION DDK(400), SDX(400), SDY(400), PD(400)
C
C
```

```
C     ASSIGN SMOOTHER PARAMETERS TO ALL GRID POINTS
C     ---------------------------------------------
      IF (ISIZE. EQ. 0) THEN     ! USE UNIFORM SMOOTHER
C
       SIGY = RYX*SIGX
       DO I=1,IMAX
       DO J=1,JMAX
       SX(I,J) = SIGX
       SY(I,J) = SIGY
       END DO
       END DO
C
      ELSE                 ! USE VARIABLE SIZE SMOOTHER
C
       DO I=1,IMAX
       DO J=1,JMAX
       CALL NEIBOR(XG(I),YG(J),SIGX,RYX,RHO,KTH,DK(I,J))
       DKB = DK(I,J)**BEXP
       SX(I,J) = CSIZE*DKB
       SY(I,J) = CSIZE*DKB*RYX
       END DO
       END DO
C
      END IF
C
C
C     CALCULATE ESTIMATE OF P AT GRID POINTS
C     --------------------------------------
      DO I=1,IMAX
      DO J=1,JMAX
      CALL KEREST(XG(I),YG(J),SX(I,J),SY(I,J),RHO,0,INTK,P(I,J))
      END DO
      END DO
C
C
      RETURN
      END
```

```
C     ================================================
      SUBROUTINE NEIBOR(XX,YY,SIGX,RYTOX,RHO,KTH,DK)
C     ================================================
C
C     FINDS DISTANCE DK TO THE KTH NEAREST NEIGHBOR TO POINT (XX,YY)
C     FROM DATA POINTS DEFINED BY (XD(I),YD(I))
C     **NEARNESS MEASURED IN MAHALANOBIS DISTANCE**
C
      COMMON/SPEC/IMAX,JMAX,NDATA
      COMMON/BASE/RESP,XD,YD
C
      DIMENSION RESP(400), YD(400), XD(400)
      DIMENSION D(400)
C
C     CALCULATE THE DISTANCE TO EACH POINT
C     ------------------------------------
      SIGY = SIGX*RYTOX
      DO I=1,NDATA
      DX = (XX - XD(I))/SIGX
      DY = (YY - YD(I))/SIGY
      D(I) = SQRT((DX**2 + DY**2 - 2.0*RHO*DX*DY)/(1.0 - RHO**2))
      END DO
C
C     SORT UNTIL KTH NEAREST NEIGHBOR FOUND
C     -------------------------------------
      DO 6 I =1,KTH
      IP1 = I + 1
      DO 5 J=IP1,NDATA
      IF (D(I).LE.D(J)) GOTO 5
      T = D(I)
      D(I) = D(J)
      D(J) = T
    5 END DO
    6 END DO
C
      DK = D(KTH)
C
      RETURN
      END
```

```
C     ================================================
      SUBROUTINE KEREST(X,Y,SIGX,SIGY,RHO,IDEL,INTK,P)
C     ================================================
C
C     SUBROUTINE FOR KERNEL REGRESSION USING THE WATSON-NADARAYA
C     FORMULATION WITH A BIVARIATE NORMAL KERNEL.
C
C     NOTE THAT THIS SUBROUTINE IS INTENDED SPECIFICALLY FOR
C     KERNEL REGRESSION USING A BINARY (0/1) RESPONSE VARIABLE (RESP).
C     HENCE THE QUANTITY P IS A PROBABILITY THAT RESP=1. HOWEVER,
C     RESP CAN ALSO BE A CONTINUOUS VARIABLE....BUT IF THE INTENDED
C     USE IS FOR CONTINUOUS RESP, CHANGES IN STMTS BEYOND LABEL 300
C     IN THIS SUBROUTINE HAVE TO BE IMPLEMENTED. ** ALSO, THE INTEGRATED
C     KERNEL ALGORITHM CANNOT BE USED FOR ORDINARY REGRESSION.**
C
C     THE INTEGRATED KERNEL OPTION IS IMPLEMENTED ASSUMING MONOTONICITY
C     AS FOLLOWS: P INCREASES AS X DECREASES AND AS Y INCREASES. THE
C     REASON WHY THE IMPLEMENTATION IS NOT MORE GENERAL IS DUE TO
C     CONSTRAINTS OF THE AVAILABLE ALGORITHM FOR INTEGRATING A BIVARIATE
C     NORMAL KERNEL.
C
C     INPUT ARGUMENTS:
C       X,Y   = X AND Y COORDINATES OF THE POINT AT WHICH THE
C               PROBABILITY/QUANTITY P IS TO BE ESTIMATED
C       SIGX  = X STD DEV (WINDOW SIZE/BANDWIDTH) OF KERNEL SMOOTHER
C       SIGY  = Y STD DEV (WINDOW SIZE/BANDWIDTH) OF KERNEL SMOOTHER
C       RHO   = CORRELATION COEFFICIENT OF KERNEL SMOOTHED
C               *** SHOULD BE SET TO ZERO FOR PRODUCT KERNEL ***
C       IDEL  = INDEX OF DATA POINT TO BE DELETED FOR ESTIMATING
C               CROSS-VALIDATED (CV) ESTIMATE OF P
C               *** SHOULD BE SET TO ZERO FOR USUAL ESTIMATE ***
C       INTK  = OPTION CODE FOR INTEGRATED KERNEL
C                 0 = USE STANDARD KERNEL
C                 1 = USE INTEGRATED KERNEL
C
C     OUTPUT ARGUMENT:
C         P   = KERNEL-SMOOTHED ESTIMATE OF PROBABILITY, OR ANY
C               OTHER QUANTITY
C
C
C     VARIABLES PASSED TO KEREST THROUGH COMMON BLOCKS
C
C       NDATA     = NUMBER OF DATA POINTS
C       RESP(K)   = RESPONSE VARIABLE VECTOR, K=1,NDATA
C       XREF(K)   = X-COORDINATES OF THE REFERENCE DATA, K=1,NDATA
C       YREF(K)   = Y-COORDINATES OF THE REFERENCE DATA, K=1,NDATA
C
      COMMON/SPEC/IMAX,JMAX,NDATA
      COMMON/BASE/RESP,XREF,YREF
C
      DIMENSION RESP(400),XREF(400),YREF(400)
      DIMENSION Z(400),R(400)
C
C
C     INITIALIZE NUMERATOR AND DENOMINATOR OF KERNEL ESTIMATOR
C     --------------------------------------------------------
      TOP = 0.
      BOT = 0.
C
      IF(INTK.EQ.1) GOTO 200
```

```
C
C  ** USING STANDARD KERNEL:
C     CALCULATE CONTRIBUTION OF EACH POINT TO THE ESTIMATE
C     OF PROBABILITY AND ACCUMULATE SUMS APPROPRIATELY
C     ================================================
C
C     CALCULATING CONSTANTS
C     ---------------------
      C1 = 2*RHO/(SIGX*SIGY)
      C2 = 0.5/(1.0 - RHO**2)
C
C
      DO 110 J=1,NDATA
C
      IF (J.EQ.IDEL) GOTO 110
C
      DX = XREF(J) - X
      DY = YREF(J) - Y
C
      ZX = DX/SIGX
      ZY = DY/SIGY
C
      Z(J) = ZX**2 + ZY**2 - C1*DX*DY
      CONTRB = EXP(-C2*Z(J))
C
      TOP = TOP + RESP(J)*CONTRB    ! NOTE CONTRB IS ONLY ADDED TO THE
      BOT = BOT + CONTRB            ! NUMERATOR WHEN Y(J)=1., NOT 0.
C
  110 END DO
C
C
      IF (BOT.GT.0.3E-38) GOTO 900  ! NO DIVISION BY ZERO PROBLEM
      GOTO 300
C
C
  200 CONTINUE
C
C     USING INTEGRATED KERNEL:
C     CALCULATE CONTRIBUTION OF EACH POINT TO THE ESTIMATE
C     OF PROBABILITY AND ACCUMULATE SUMS APPROPRIATELY
C     ================================================
C
      DO 210 J=1,NDATA
C
      IF(J.EQ.IDEL) GOTO 210
C
      DX = XREF(J) - X
      DY = YREF(J) - Y
C
      ZX = DX/SIGX
      ZY = DY/SIGY
C
C     SET MAX LIMITS ON ZX AND ZY TO AVOID PROBLEMS WITH
C     SUBROUTINE MDBNOR (THERE IS A BUG IN THE SUBROUTINE)
C     ----------------------------------------------------
      IF (ZX.LT.-4.0) ZX = -4.0
      IF (ZX.GT. 4.0) ZX =  4.0
      IF (ZY.LT.-4.0) ZY = -4.0
      IF (ZY.GT. 4.0) ZY =  4.0
C
```

```
      CALL MDBNOR(ZX,ZY,RHO,Q,IER)
      CALL MDNOR(ZX,Q1)
      CALL MDNOR(ZY,Q0)
C
      CONTR1 = RESP(J)*(Q1 - Q)
      CONTR0 = (1.0 - RESP(J))*(Q0 - Q)
C
      TOP = TOP + CONTR1
      BOT = BOT + CONTR1 + CONTR0
C
  210 END DO
C
C
      IF (BOT.GT.0.3E-38) GOTO 900  ! NO DIVISION BY ZERO PROBLEM
C
  300 CONTINUE
C
C     FIX DIVISION BY ZERO PROBLEM
C     ============================
C
      IF (INTK.EQ.0) GOTO 310
C
C     CALCULATING DISTANCE MEASURE
C     ----------------------------
      C1 = 2*RHO/(SIGX*SIGY)
      DO J=1,NDATA
      DX = X - XREF(J)
      DY = Y - YREF(J)
      ZX = DX/SIGX
      ZY = DY/SIGY
      Z(J) = ZX**2 + ZY**2 - C1*DX*DY
      END DO
C
  310 CONTINUE
C
C     CREATE R-VECTOR & SORT TO FIND THE FOUR NEAREST NEIGHBORS
C     ---------------------------------------------------------
      DO I= 1,NDATA
      R(I) = RESP(I)
      END DO
C
      DO 6 I=1,4
      IP1 = I + 1
      DO 5 J= IP1,NDATA
      IF (Z(I).LE.Z(J)) GOTO 5
      T = Z(I)
      Z(I) = Z(J)
      Z(J) = T
      TR = R(I)
      R(I) = R(J)
      R(J) = TR
    5 END DO
    6 END DO
C
C     THEN FIND THE SUM OF THE TOP FOUR R(I)
C     --------------------------------------
      SUM = 0.
      DO I=1,4
      SUM = SUM + RESP(I)
      END DO
```

```
C
C     SET P DEPENDING IF (X,Y) IN  RESP = 1  OR  0  DOMAIN
C     ----------------------------------------------------
      IF (SUM.LE.2.0) THEN
       P = 0.0000001
      ELSE
       P = 0.9999999
      END IF
C
      GOTO 999
C
  900 CONTINUE
C
C     P ESTIMATE WITH NONZERO DIVISOR
C     -------------------------------
      P = TOP/BOT
C
      IF (P.LE.0.0000001) P = 0.0000001
      IF (P.GE.0.9999999) P = 0.9999999
C
  999 CONTINUE
      RETURN
      END
```

```
C     ================================================================
      SUBROUTINE SLBOX(LOGX,XLO,XHI,NXDIV,LOGY,YLO,YHI,NYDIV,
     1                                          XLAB,YLAB,PLAB,IOPTC)
C     ================================================================
C
C                                                    S.L. JUNE 29,1984
C
C     INITIALIZES PLOTTER, DRAWS A BOX WITH AXES AND LABELS
C
C     USAGE: CALL SLBOX(LOGX,XLO,....,IOPTC,MARGIN(UL,UR,VB,VT))
C
C     ARGUMENTS:
C         LOGX = LOG OPTION FOR THE X-SCALE
C          XLO = LOWER BOUND VALUE OF X-AXIS
C          XHI = UPPER BOUND VALUE OF X-AXIS
C        NXDIV = NUMBER OF SCALE DIVISIONS OF X-AXIS
C         LOGY = LOG OPTION FOR THE Y-SCALE
C          YLO = LOWER BOUND VALUE OF Y-AXIS
C          YHI = UPPER BOUND VALUE OF Y-AXIS
C        NYDIV = NUMBER OF SCALE DIVISIONS OF Y-AXIS
C         XLAB = X-AXIS LABEL --- CHARACTER*(*)
C         YLAB = Y-AXIS LABEL --- CHARACTER*(*)
C         PLAB = PLOT LABEL ----- CHARACTER*(*)
C        IOPTC = OPTION CONTROL INDICATOR
C                  0 = USE DEFAULT OPTIONS
C                  1 = USE OPTIONS SPECIFIED BY ENTRY MARGIN:
C
C                 MARGIN(UL,UR,...) IS OPTIONAL PART OF CALL AND IS
C                 ONLY NECESSARY IF IOPTC = 1
C                   ARGUMENTS:
C                    UL = LEFT MARGIN  (IN INCHES)
C                    UR = RIGHT MARGIN (IN INCHES)
C                    VB = BOTTOM MARGIN (IN INCHES)
C                    VT = TOP MARGIN (IN INCHES)
C                NOTE: MARGINS ARE FROM THE PLOTTING AREA LIMITS,WHICH
C                ON THE IMAGEN PLOTTER HAS A WIDTH OF 10.23 INCHES
C                AND A HEIGHT OF 7.5 INCHES.
C
C     NOTE: *WHEN USING LOG SCALE OPTIONS LOGX AND/OR LOGY HAVE
C            THE FOLLOWING VALUES:
C                  0 = NORMAL SCALE
C                  1 = LOG SCALE
C           *XLO AND YLO MUST CORRESPOND TO THE POWER EXPONENTS OF
C            10 WHICH BOUND THE LOWER VALUE OF THE PLOTTED POINTS
C           *XHI AND YHI MUST CORRESPOND TO THE POWER EXPONENTS OF
C            10 WHICH BOUND THE  UPPER VALUES OF THE PLOTTED POINTS
C           *NXDIV AND NYDIV MUST EQUAL THE NUMBER OF LOG CYCLES BETWEEN
C            XLO AND XHI  OR  YLO AND YHI
C
C
C     NOTE: IN THIS SUBROUTINE, VARIABLES WITH FIRST LETTER X OR Y
C           DENOTE USER COORDINATES. VARIABLES WITH FIRST LETTER
C           U OR V DENOTE ACTUAL PHYSICAL COORDINATES (IN INCHES)
C           OF PLOTS ON THE IMAGEN.
C
C
      CHARACTER*(*) XLAB,YLAB,PLAB
C
C
      CALL T4025
```

```
      CALL LOCATE(0.,100.,0.,100.)
C
C     BRANCH TO STMT 20 IF DEFAULT MARGINS NOT DESIRED,
C       BUT SPECIFIED THROUGH ENTRY MARGIN
C
      IF(IOPTC.NE.0) GOTO 20
C
C     SET X-AXIS LEFT AND RIGHT DEFAULT MARGINS (IN INCHES)
C
      UMARL = 2.23
      UMARR = 2.0
C
C     SET Y-AXIS BOTTOM AND TOP DEFAULT MARGINS (IN INCHES)
C
      VMARB = 1.5
      VMART = 2.0
C
C
   20 CONTINUE
C
C
C     SET RELATIVE X-POSITION FOR Y-LABEL (IN INCHES)
C
      UYLAB = -0.8
C
C     SET RELATIVE Y-POSITION FOR X-LABEL (IN INCHES)
C
      VXLAB = -0.5
C
C     SET RELATIVE POSITION FOR PLOT LABEL
C
      VPLAB = -0.8
C
C     CONVERT PHYSICAL COORDINATES TO USER COORDNATES
C
C     CALCULATE  WIDTH AND HEIGHT OF PLOTTING AREA (IN INCHES)
C         ON THE IMAGEN PLOTTER
C
      ULEN = 10.23 - (UMARL + UMARR)
      VLEN = 7.5 - (VMARB + VMART)
C
C
      XLEN = XHI - XLO
      DX = XLEN/FLOAT(NXDIV)
      XLIML = XLO - XLEN * UMARL/ULEN
      XLIMR = XHI + XLEN * UMARR/ULEN
C
      YLEN = YHI - YLO
      DY = YLEN/FLOAT(NYDIV)
      YLIMB = YLO - YLEN * VMARB/VLEN
      YLIMT = YHI + YLEN * VMART/VLEN
C
      YXLAB = YLO + VXLAB * YLEN/VLEN
      YPLAB = YLO + VPLAB * YLEN/VLEN
      XYLAB = XLO + UYLAB * XLEN/ULEN
C
C     SETTING PLOTTER LIMITS
C
      CALL AREA(XLIML,XLIMR,YLIMB,YLIMT)
C
```

```
C     PLOTTING AXES AND ANNOTATION
C
      CALL LETTER(1.6,0.7,0.,0.)
      CALL TSAXE(0,XLO,YLO,DX,NXDIV,0,LOGX,0,0)
      CALL TSAXE(1,XHI,YLO,DY,NYDIV,0,LOGY,0,0)
      CALL TSAXE(2,XHI,YHI,DX,NXDIV,0,LOGX,0,0)
      CALL LETTER(1.6,0.7,0.,0.)
      CALL TSAXE(3,XLO,YHI,DY,NYDIV,0,LOGY,0,0)
C
C     WRITING AXES LABELS
C
      XMIDPT = 0.5*(XLO + XHI)
      YMIDPT = 0.5*(YLO + YHI)
C
      CALL LORG(6)
      CALL LETTER(2.0,0.7,0.,0.)
      CALL MOVE(XMIDPT,YXLAB)
      CALL LABEL(XLAB)
C
      CALL LETTER(2.8,0.7,0.,0.)
      CALL MOVE(XMIDPT,YPLAB)
      CALL LABEL(PLAB)
C
      CALL LORG(4)
      CALL LETTER(2.0,0.7,90.,0.)
      CALL MOVE(XYLAB,YMIDPT)
      CALL LABEL(YLAB)
C
C
      RETURN
C
C
C     OPTIONAL ENTRY OF MARGIN LIMITS IN INCHES
C     -----------------------------------------
      ENTRY MARGIN(UL,UR,VB,VT)
C
      UMARL = UL
      UMARR = UR
      VMARB = VB
      VMART = VT
C
      RETURN
C
C
      END
```

```
C     ========================
      SUBROUTINE SLTOUR(VINT)
C     ========================
C
C                                          S.L. JUNE 26,1984
C                                          S.L. FEB  14,1985
C
C     SETS UP THE GRID OF DATA CONTAINED IN MATRIX Q FOR
C     SUBROUTINE CONPLT AND CALLS CONPLT TO PRODUCE A CONTOUR
C     PLOT OF THE DATA
C
C     **IT IS ASSUMED THAT THE USER HAS ALREADY INITIALIZED
C       AND DEFINED THE PLOTTING LIMITS (USING PENPLOT CALLS
C       OR BY OTHER SUBROUTINES, E.G., SLBOX) BEFORE CALLING
C       THIS SUBROUTINE.
C
C     INPUT VARIABLES:
C       Q(I,J)    = MATRIX OF VALUES TO BE CONTOURED
C       X(I)      = VECTOR OF X-COORDINATES OF THE GRID POINTS
C       Y(J)      = VECTOR OF Y-COORDINATES OF THE GRID POINTS
C       IMAX      = MAX NUMBER OF X-COORDINATE POINTS
C       JMAX      = MAX NUMBER OF Y-COORDINATE POINTS
C       VINT      = CONTOUR INTERVAL SPACING (BETWEEN VALUES IN Q)
C
C
C     NOTE ON VARIABLE XX:
C
C       XX(3,5) IS THE INPUT FORMAT OF THE DATA REQUIRED BY CONPLT.
C
C        *THERE ARE 4 POINTS FOR EACH RECTANGULAR CELL OF THE GRID.
C         THE CENTER POINT OF THE CELL IS DETERMINED AS A 5th POINT,
C         BY CONPLT. HENCE THE DIMENSION 5. SEE DIAGRAM BELOW.
C
C        *XX(1,I) IS THE X-COORDINATE OF THE Ith POINT
C         XX(2,I) IS THE Y-COORDINATE OF THE Ith POINT
C         XX(3,I) IS THE VALUE TO BE CONTOURED AT TH Ith POINT
C
C
      COMMON/SPEC/IMAX,JMAX,NDATA
      COMMON/GRID/X,Y,Q
C
      DIMENSION Q(50,50),X(50),Y(50),XX(3,5),M(4)
C
      DATA M/1,2,3,4/    ! NOTE: M IS THE VECTOR OF NODAL POINT NUMBERS
C                        ! BUT IN THIS PROGRAM, SUCH A NUMBERING SYSTEM
C                        ! IS NOT USED
C
      NXCELL = IMAX - 1
      NYCELL = JMAX - 1
C
      DO J=1,NYCELL
      DO I=1,NXCELL
C
      XX(1,1) = X(I)
      XX(2,1) = Y(J)
      XX(3,1) = Q(I,J)
C
      XX(1,2) = X(I+1)
      XX(2,2) = Y(J)
      XX(3,2) = Q(I+1,J)
C
      XX(1,3) = X(I+1)
      XX(2,3) = Y(J+1)
      XX(3,3) = Q(I+1,J+1)
```

```
!   I,J+1                    I+1,J+1
!     4                        3
!     *------------------------*
!     !                        !
!     !           *5           !
!     !                        !
!     *------------------------*
!     1                        2
!    I,J                     I+1,J
```

```
C
      XX(1,4) = X(I)
      XX(2,4) = Y(J+1)
      XX(3,4) = Q(I,J+1)
C
C     FOR EACH RECTANGULAR CELL, DRAW THE CONTOURS
C         ----
      CALL CONPLT(XX,M,VINT)
C
      END DO
      END DO
C
      RETURN
      END
C
C     ================================
      SUBROUTINE CONPLT(XX,M,VINT)
C     ================================
C
C     PLOTS CONTOURS WITHIN A QUDRILATERAL OR TRIANGULAR ELEMENT
C
      DIMENSION XX(3,5),M(4),II(3)
      NUM = 1
      NML = M(3)
      IF (M(4).EQ.0) GOTO 50
C
C-----FIND THE VALUE OF THE CENTER POINT OF THE QUADRILATERAL
C
      NML = 5
      NUM = 4
      XX(1,NML) = 0.0
      XX(2,NML) = 0.0
      XX(3,NML) = 0.0
      DO 10 I=1,4
      XX(1,NML) = XX(1,NML) + XX(1,M(I))*0.25
      XX(2,NML) = XX(2,NML) + XX(2,M(I))*0.25
      XX(3,NML) = XX(3,NML) + XX(3,M(I))*0.25
   10 CONTINUE
C
C-----LOOP OVER THE SUBTRIANGLES
C
   50 CONTINUE
      DO 100 I=1,NUM
      II(1) = M(I)
      II(2) = M(I+1)
      IF (I.EQ.4) II(2) = M(1)
      II(3) = NML
C
C-----ORDER THE POINTS ACCORDING TO ASCENDING VALUES OF XX(3,*)
C
      DO 60 J=1,2
      DO 60 L=J,3
      IF (XX(3,II(J)).GE.XX(3,II(L))) GOTO 60
      ITE = II(J)
      II(J) = II(L)
      II(L) = ITE
```

```
   60 CONTINUE
C
C-----CALCULATE THE NUMBER OF CONTOUR LINES IN THIS TRIANGLE
C
      IF(XX(3,II(1)).EQ.XX(3,II(3))) GOTO 100
      IMAX = XX(3,II(1))/VINT
      IF (XX(3,II(1)).LT.0.0) IMAX = IMAX - 1
      VMAX = FLOAT(IMAX)*VINT
      VDIF = VMAX - XX(3,II(3))
      NLIN = (VDIF/VINT) + 1
      IF (NLIN.LE.0) GOTO 100
C
C-----DRAW THE LINES
C
      I1 = 1
      DO 80 L=1,NLIN
      IF (VMAX.GT.0.0) CALL LINE(0)
      IF (VMAX.EQ.0.0) CALL LINE(6)
      IF (VMAX.LT.0.0) CALL LINE(7)
      IF (VMAX.LT.XX(3,II(2))) I1 = 2
      RAT = (XX(3,II(1)) - VMAX)/(XX(3,II(1)) - XX(3,II(3)))
      XP = XX(1,II(1))*(1.0 - RAT) + XX(1,II(3))*RAT
      YP = XX(2,II(1))*(1.0 - RAT) + XX(2,II(3))*RAT
      CALL MOVE(XP,YP)
      RAT = XX(3,II(I1)) - XX(3,II(I1+1))
      IF (RAT.EQ.0.0) GOTO 75
      RAT = (XX(3,II(I1)) - VMAX)/RAT
      XP = XX(1,II(I1))*(1.0 - RAT) + XX(1,II(I1+1))*RAT
      YP = XX(2,II(I1))*(1.0 - RAT) + XX(2,II(I1+1))*RAT
      CALL DRAW(XP,YP)
   75 VMAX = VMAX - VINT
   80 CONTINUE
  100 CONTINUE
C
C
      RETURN
      END
```

```
C     ==========================
      SUBROUTINE DATPLT(ISYM)
C     ==========================
C
C     PLOTS THE DATA POINTS
C
C     ISYM  =  SYMBOL CONTROL POINT
C                0 = DISTINGUISH BETWEEN RESP = 1 OR 0
C                1 = PLOT ALL POINTS THE SAME
C
      COMMON/SPEC/IMAX,JMAX,NDATA
      COMMON/BASE/RESP,XD,YD
C
      DIMENSION  RESP(400), XD(400), YD(400)
C
      CALL LORG(5)
C
      IF (ISYM.EQ.1) GOTO 100
C
      DO K=1,NDATA
C
      CALL MOVE(XD(K),YD(K))
      IRESP = NINT(RESP(K))
      IF (IRESP.EQ.1) THEN
       CALL LETTER(2.0, 0.7, 0.0, 0.0)
       CALL LABEL('+')
      ELSE
       CALL LETTER(1.0,1.25, 0.0, 0.0)
       CALL LABEL('O')
      END IF
C
      END DO
      GOTO 999
C
  100 CONTINUE
C
      CALL LETTER(2.0,0.7,0.0,0.0)
      DO  K=1,NDATA
      CALL MOVE(XD(K),YD(K))
      CALL LABEL('+')
      END DO
C
  999 CONTINUE
      RETURN
      END
```